图灵数学·统计学丛书

普林斯顿微积分读本（修订版）

[美] Adrian Banner ◎著　　杨爽 赵晓婷 高璞 ◎译

The Calculus Lifesaver

All the Tools You Need to Excel at Calculus




中国工信出版集团


人民邮电出版社
POSTS & TELECOM PRESS



Adrian Banner
（阿德里安·班纳）

澳大利亚新南威尔士大学数学学士及硕士，普林斯顿大学数学博士。2002年起任职于INTECH公司，现为INTECH公司首席执行官兼首席投资官。同时，他在普林斯顿大学数学系任兼职教师。

配套课程视频：
http://www.calclifesaver.com/

数字版权声明

图灵社区的电子书没有采用专有客户端，您可以在任意设备上，用自己喜欢的浏览器和PDF阅读器进行阅读。

但您购买的电子书仅供您个人使用，未经授权，不得进行传播。

我们愿意相信读者具有这样的良知和觉悟，与我们共同保护知识产权。

如果购买者有侵权行为，我们可能对该用户实施包括但不限于关闭该帐号等维权措施，并可能追究法律责任。

TURING 图灵数学 · 统计学丛书

普林斯顿微积分读本（修订版）

[美] Adrian Banner ◎著　杨爽 赵晓婷 高璞 ◎译

The Calculus Lifesaver

All the Tools You Need to Excel at Calculus



人民邮电出版社

北京

图书在版编目（CIP）数据

普林斯顿微积分读本 / (美) 阿德里安•班纳著 ;
杨爽, 赵晓婷, 高璞译. -- 2版. -- 北京 : 人民邮电出
版社, 2016.10（2019.11重印）
（图灵数学•统计学丛书）
书名原文: The Calculus Lifesaver: All the
Tools You Need to Excel at Calculus
ISBN 978-7-115-43559-0

Ⅰ. ①普… Ⅱ. ①阿… ②杨… ③赵… ④高… Ⅲ.
①微积分 Ⅳ. ①O172

中国版本图书馆CIP数据核字(2016)第216695号

内 容 提 要

本书阐述了求解微积分的技巧，详细讲解了微积分基础、极限、连续、微分、导数的应用、积分、无穷级数、泰勒级数与幂级数等内容，旨在教会读者如何思考问题从而找到解题所需的知识点，着重训练大家自己解答问题的能力。

本书适用于大学低年级学生、高中高年级学生、想学习微积分的数学爱好者以及广大数学教师，既可作为教材、习题集，也可作为学习指南，同时还有利于教师备课。

◆ 著　　[美] Adrian Banner
译　　杨　爽　赵晓婷　高　璞
责任编辑　傅志红
责任印制　彭志环
◆ 人民邮电出版社出版发行　北京市丰台区成寿寺路11号
邮编　100164　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
固安县铭成印刷有限公司印刷
◆ 开本：700×1000　1/16
印张：41.75
字数：895千字　2016年10月第2版
印数：92 001 – 95 500册　2019年11月河北第24次印刷
著作权合同登记号　图字：01-2009-3812号

定价：99.00元

读者服务热线：(010)51095183转600　印装质量热线：(010)81055316

反盗版热线：(010)81055315

广告经营许可证：京东工商广登字 20170147 号

版 权 声 明

献给亚里

译　者　序

对于大多数学生来说, 微积分或许是他们曾经上过的倍感迷茫且最受挫折的一门课程了. 而本书, 不仅让学生能有效地学习微积分, 更重要的是提供了战胜微积分的必备工具.

本书源于风靡美国普林斯顿大学的阿德里安 · 班纳的微积分复习课程. 他激励了一些考试前想获得优秀但考试结果却平平的学生.

对于任何单变量微积分的课程, 本书既可以作为教科书, 也可以用作学习指南, 对于全英文授课的教师来说更是一个得力助手. 作者班纳是美国普林斯顿大学的著名数学教授并担任新技术研究中心主任. 班纳教授的授课风格是非正式、有吸引力并完全不强求的, 甚至在不失其详尽性的基础上又增添了许多娱乐性, 而且他不会跳过讨论一个问题的任何步骤.

作者独创的"内心独白"方式, 即写出问题求解过程中学生们应遵循的思考过程, 为我们提供了不可或缺的推理过程以及求解方案. 本书的重点在于培养问题求解的能力, 其中涉及的例题从简单到复杂并对微积分理论进行了深入探讨. 读者会在非正式的对话语境中体会到微积分的无穷魅力.

本书特点:

- 可作为任何单变量微积分教科书的学习指南;
- 非正式的、娱乐性的且非强求的对话语境风格;
- 丰富的在线视频;
- 大量精选例题 (从简单到复杂) 提供了一步一步的推理过程;
- 定理和方法有证明, 还有诸多实际应用;
- 详细探讨了诸如无穷级数这样的难点问题.

这样的一本经典著作将易用性与可读性以及内容的深度与数学的严谨完美地结合在一起. 对于每一个想要掌握微积分的学生来说, 本书都是极好的资源. 当然, 非数学专业的学生也将大大受益.

在翻译本书的过程中, 译者虽然尽最大努力尊重原文, 并尽可能避免直译产生的歧义, 但是由于才疏学浅, 难免存在翻译不当之处, 敬请广大读者批评指正, 以便再版时更正.

本书能得以顺利出版, 首先要感谢人民邮电出版社图灵公司的大力支持; 同时, 首都经济贸易大学华侨学院信管系的全体教师也给予了无私的帮助, 在此一并表示

衷心感谢. 最后感谢我的家人在本书翻译过程中所给予的支持与鼓励, 尤其是爱女芮绮!

《普林斯顿微积分读本》
微笑着面对数学的世界
积累着超越无穷的力量
分化出化解疑难的翅膀
求解出优化问题的阳光
生成了数学天空的晴朗
秘籍 —— 放飞自己的理想

谨以此诗献给爱女芮绮以及喜爱数学的新生代!

杨 爽
首都经济贸易大学华侨学院信管系

前　　言

本书旨在帮助你学习单变量微积分的主要概念, 同时也致力于教会你求解问题的技巧. 无论你是第一次接触微积分, 还是为了准备一次测验, 或是已经学过微积分还想再温习一遍, 我都希望本书能够对你有所帮助.

写作本书的灵感来自我在普林斯顿大学的学生们. 他们在过去的几年里发现, 与课堂授课、作业讲解以及他们的教科书一样, 本书的初稿是很有帮助的学习指南. 以下是他们在学习过程中提出的一些你可能也想问的问题.

这本书为什么这么厚? 我是假设你真的想要掌握这门课程, 而不只是想囫囵吞枣, 一知半解, 所以你已经准备好投入一些时间和精力, 去阅读并理解这些详尽的阐述.

阅读之前, 我需要知道些什么? 你需要了解一些基本的代数知识, 并且要知道如何求解简单的方程式. 本书的前两章涵盖了你所需要的大部分的微积分预备知识.

啊! 下周就要期末考试了, 我还什么都不知道呢! 从哪里开始啊? 接下来的几页就会介绍如何使用本书来备考.

例题的求解过程在哪里? 我所看到的只是大量的文字与少量的公式. 首先, 看一个求解过程并不能教会你应该怎样思考. 所以我通常试图给出一种 “内心独白”, 即当你尝试求解问题的时候, 脑海中应该经历怎样的思考过程. 最后, 你想到了求解问题的所有知识点, 但仍然需要用正确的方式把它们全部写出来. 我的建议是, 先看懂并理解问题的求解方法, 然后再返回来尝试自己解答.

定理的证明哪儿去了? 本书中的大部分定理都以某种方式被验证了. 在附录 A 中可以找到更多正式的证明过程.

主题没有次序! 我该怎么办呢? 学习微积分没有什么标准次序. 我选择的顺序是有效的, 但你可能还得通过搜索目录来查找你需要的主题, 其余的可以先忽略. 我也可能遗漏了一些主题. 为什么不尝试给我发送电子邮件呢? 地址是 adrian@calc-lifesaver.com. 你一定想不到, 我可能会为你写一个附加章节 (也为下一版写, 如果有的话!).

你使用的一些方法和我学到的不一样. 到底谁的正确, 我的任课老师的还是你的? 希望我们都没错! 如果还有疑问, 就请教你的任课老师什么是对的吧.

页边空白处怎么没有微积分的历史和有趣的史实呢? 本书中有一点微积分历史内容, 但不在这里过多分散我们的注意力. 如果你想记下这些历史内容, 就请

阅读一本关于微积分历史的书[①]吧, 那才更有趣, 而且比零零散散的几句话更值得关注.

我们学校可以用这本书作为教材吗? 这本书配有很好的习题集, 可以作为一本教材, 也可以用作一本学习指南. 你的任课老师也会发现这本书很有助于备课, 特别是在问题求解的技巧方面.

这些录像是什么? 在网站 www.calclifesaver.com 上, 你可以找到我过去复习课的录像, 其中涉及了很多 (但不是全部) 本书的章节和例题.

如何使用这本书备考

如果你快要参加考试了, 那么发挥本书效用的机会就来了. 我很同情你的处境, 因为你没有时间阅读整本书的内容! 但是你不用担心, 后面的那张表会标出本书的主要章节, 来帮助你备考. 此外, 纵观整本书, 下列图标会出现在书中页边空白处, 让你快速识别什么是重要内容.



- 例题求解过程始于此行.



- 这里非常重要.



- 你应当自己尝试解答本题.



- 注意: 这部分内容大多是为感兴趣的读者准备的. 如果时间有限, 就请跳到下一节.

此外, 一些重要的公式或定理带有边框, 一定要好好学啊.

两个通用的学习小贴士

- 把你自己总结的所有重要的知识点和公式都写出来, 以便记忆. 虽说数学不是死记硬背, 但也有一些关键的公式和方法, 最好是你能自己写得出来. 好记性不如烂笔头嘛! 通常来说, 做总结足以巩固和加强你对所学知识的理解. 这也是我没有在每一章的结尾部分做要点总结的主要原因. 如果你自己去做, 那将会更有价值.
- 尝试自己做一些类似的考试题, 比如你们学校以前的期末试题, 并在恰当的条件下进行测验. 这将意味着遵守不间断, 不吃饭, 不看书, 不打手机, 不发电子邮件, 不发信息等诸如此类的考试规则. 完成之后, 再看看你是否可以得到一套标准答案来评阅试卷, 或请人帮你评阅.

① 对微积分历史感兴趣的读者, 可参阅《微积分的历程: 从牛顿到勒贝格》(人民邮电出版社, 2010). ——编者注

考试复习的重要章节 (按主题划分)

主题	子主题	节
微积分基础	直线	1.5
	其他常用图像	1.6
	三角学基础	2.1
	$[0,\pi/2]$ 以外的三角函数	2.2
	三角函数的图像	2.3
	三角恒等式	2.4
	指数函数与对数函数	9.1
极限	三明治定理	3.6
	多项式的极限	第 4 章全部
	导数伪装的极限	6.5
	三角函数的极限	7.1(跳过 7.1.5)
	指数函数与对数函数的极限	9.4
	洛必达法则	14.1
	极限问题的总结	14.2
连续性	定义	5.1
	介值定理	5.1.4
微分	定义	6.1
	求导法则 (例如, 乘积法则/商法则/链式求导法则)	6.2
	求切线方程	6.3
	分段函数的导数	6.6
	画导函数图像	6.7
	三角函数的导数	7.2, 7.2.1
	隐函数求导	8.1
	指数函数与对数函数求导	9.3
	取对数求导法	9.5
	双曲函数	9.7
	反函数	10.1
	反三角函数	10.2
	反双曲函数	10.3
	求导定积分	17.5

(续)

主题	子主题	节
导数的应用	相关变化率	8.2
	指数增长与指数衰变	9.6
	求全局最大值与全局最小值	11.1.3
	罗尔定理/中值定理	11.2, 11.3
	临界点的分类	11.5, 12.1.1
	求拐点	11.4, 12.1.2
	画图	12.2, 12.3
	最优化	13.1
	线性化/微分	13.2
	牛顿法	13.3
积分	定义	16.2(跳过 16.2.1)
	基本性质	16.3
	求面积	16.4
	估算积分	16.5, 附录 B
	平均值/中值定理	16.6
	基本例子	17.4, 17.6
	换元法	18.1
	分部积分法	18.2
	部分分式	18.3
	三角函数的积分	19.1, 19.2
	三角换元法	19.3(跳过 19.3.6)
	积分技巧的总结	19.4
运动	速度与加速度	6.4
	负常数加速度	6.4.1
	简谐运动	7.2.2
	求位移	16.1.1
反常积分	基本知识	20.1, 20.2
	求解技巧	第 21 章全部
无穷级数	基本知识	22.1.2, 22.2
	求解技巧	第 23 章全部
泰勒级数与幂级数	估算和误差估算	第 25 章全部
	幂级数/泰勒级数问题	第 26 章全部

(续)

主题	子主题	节
微分方程	可分一阶	30.2
	一阶线性	30.3
	常系数	30.4
	建模	30.5
其他话题	参数方程	27.1
	极坐标	27.2
	复数	28.1 ~ 28.5
	体积	29.1, 29.2
	弧长	29.3
	表面积	29.4

除非特殊说明, 标明 “节” 的一栏包括其下所有小节. 例如, 6.2 节包括从 6.2.1 到 6.2.7 的所有小节.

致　　谢

感谢所有在我写作本书过程中给予我支持和帮助的人. 我的学生们长久以来在给我教益、喜悦和快乐, 他们的意见使我受益匪浅. 特别感谢我的编辑 Vickie Kearn、制作编辑 Linny Schenck 和设计师 Lorraine Doneker, 感谢他们对我的所有帮助和支持, 还要感谢 Gerald Folland, 他的很多真知灼见对本书的改善有很大的贡献. 此外, 感谢 Ed Nelson、Maria Klawe、Christine Miranda、Lior Braunstein、Emily Sands、Jamaal Clue、Alison Ralph、Marcher Thompson、Ioannis Avramides、Kristen Molloy、Dave Uppal、Nwanneka Onvekwusi、Ellen Zuckerman、Charles MacCluer 和 Gary Slezak, 本书中的很多修正都得益于他们的意见和建议.

感谢下列普林斯顿大学数学系的教员和工作人员对我的大力支持：Eli Stein、Simon Kochen、Matthew Ferszt 和 Cott Kenny. 我也要感谢我在 INTECH 的同事们给予的支持, 特别是 Bob Fernholz、Camm Maguire、Marie D'Albero 和 Vassilios Papathanakos, 他们提出了一些优秀的审读建议. 我还要感谢我高二、高三的数学老师 ——William Pender, 他绝对是世界上最好的微积分老师. 这本书中很多方法都是从他的教学中获得了启发. 我希望他能原谅我曲线不画箭头, 所有的坐标轴上没有标注, 以及在每一个 $+C$ 后都没有写 "对于任意一个常数 C".

我的朋友和家人都给了我无私的支持, 尤其是我的父母 Freda 和 Michael、姐姐 Carly、祖母 Rena, 还有姻亲 Marianna 和 Michael. 最后, 我要特别感谢我的妻子 Amy 在我写书过程中对我的帮助和理解, 她总是陪伴在我身边. (还要感谢她为我画的 "爬山者图标".)

目　　录

第 1 章　函数、图像和直线

不借助函数却想去做微积分, 这无疑会是你所能做的最无意义的事情之一. 如果微积分也有其营养成分表, 那么函数肯定会排在最前面, 而且是占一定优势. 因此, 本书的前两章旨在让你温习函数的主要性质. 本章包含对下列主题的回顾:

- 函数, 其定义域、上域、值域和垂线检验;
- 反函数和水平线检验;
- 函数的复合;
- 奇函数与偶函数;
- 线性函数和多项式的图像, 以及对有理函数、指数函数和对数函数图像的简单回顾;
- 如何处理绝对值.

下一章会涉及三角函数. 好啦, 就让我们开始吧, 一起来回顾一下到底什么是函数.

1.1　函　数

函数是将一个对象转化为另一个对象的规则. 起始对象称为**输入**, 来自称为**定义域**的集合. 返回对象称为**输出**, 来自称为**上域**的集合.

来看一些函数的例子吧.

- 假设你写出 $f(x) = x^2$, 这就定义了一个函数 f, 它会将任何数变为自己的平方. 由于你没有说明其定义域或上域, 我们不妨假设它们都属于 $\mathbb{R}$, 即所有实数的集合. 这样, 你就可以将任何实数平方, 并得到一个实数. 例如, f 将 2 变为 4、将 $-1/2$ 变为 $1/4$, 将 1 变为 1. 最后一个变换根本没有什么变化, 但这没问题, 因为转变后的对象不需要有别于原始对象. 当你写出 $f(2) = 4$ 的时候, 这实际上意味着 f 将 2 变为 4. 顺便要说的是, f 是一个变换规则, 而 $f(x)$ 是把这个变换规则应用于变量 x 后得到的结果. 因此, 说 "$f(x)$ 是一个函数" 是不正确的, 应该说 "f 是一个函数".
- 现在, 令 $g(x) = x^2$, 其定义域仅包含大于或等于零的数 (这样的数称为**非负的**). 它看上去好像和函数 f 是一样的, 但它们实际不同, 因为各自的定义域不同. 例如, $f(-1/2) = 1/4$, 但 $g(-1/2)$ 却是没有定义的. 函数 g 会拒绝非其定义域中的一切. 由于 g 和 f 有相同的规则, 但 g 的定义域小于 f 的定义域, 因而我们说 g 是由**限制** f 的定义域产生的.

- 仍然令 $f(x) = x^2$, $f(\text{马})$ 会是什么呢？这显然是无定义的, 因为你不能平方一匹马呀. 另一方面, 让我们指定 "$h(x) = x$ 的腿的数目", 其中 h 的定义域是所有动物的集合. 这样一来, 我们就会得到 $h(\text{马}) = 4$, $h(\text{蚂蚁}) = 6$, $h(\text{鲑鱼}) = 0$. 因为动物腿的数目不会是负数或者分数, 所以 h 的上域可以是所有非负整数的集合. 顺便问一下, $h(2)$ 会是什么呢？当然, 这也是没有定义的, 因为 2 不在 h 的定义域中. "2" 究竟会有几条腿呢？这个问题实际上没有任何意义. 你或许也可以认为 $h(\text{椅子}) = 4$, 因为多数椅子都有四条腿, 但这也没有意义, 因为椅子不是动物, 所以 "椅子" 不在 h 的定义域中. 也就是说, $h(\text{椅子})$ 是没有定义的.
- 假设你有一条狗, 它叫 Junkster. 可怜的 Junkster 不幸患有消化不良症. 它吃点东西, 嚼一会儿, 试图消化食物, 可每次都失败, 都会吐出来. Junkster 将食物变成了 …… 我们可以令 "$j(x) =$ Junkster 吃 x 时呕吐物的颜色", 其中 j 的定义域是 Junkster 所吃的食物的集合, 其上域是所有颜色的集合. 为了使之有效, 我们必须认为如果 Junkster 吃了玉米面卷, 它的呕吐物始终是一种颜色 (假设是红色的吧). 如果有时候是红色的, 而有时候是绿色的, 那就不太好了. **一个函数必须给每一个有效的输入指定唯一的输出.**

现在我们要来看看函数**值域**的概念. 值域是所有可能的输出所组成的集合. 你可以认为函数转变其定义域中的一切, 每次转变一个对象; 转变后的对象所组成的集合称作值域. 可能会有重复, 但这也没什么.

那么, 为什么值域和上域不是一回事呢？值域实际上是上域的一个子集. 上域是**可能**输出的集合, 而值域则是**实际**输出的集合. 下面给出上述函数的值域.

- 如果 $f(x) = x^2$, 其定义域和上域均为 $\mathbb{R}$, 那么其值域是非负数的集合. 毕竟, 平方一个数, 其结果不可能是负数. 那你又如何知道值域是**所有**的非负数呢？其实, 如果平方每一个数, 结果一定包括所有的非负数. 例如, 平方 $\sqrt{2}$ (或 $-\sqrt{2}$), 结果都是 2.
- 如果 $g(x) = x^2$, 其定义域仅为非负数, 但其上域仍是所有实数 $\mathbb{R}$, 那么其值域还是非负数的集合. 当平方每一个非负数时, 结果仍然包括所有的非负数.
- 如果 $h(x)$ 是动物 x 的腿的数目, 那么其值域就是**任何**动物可能会有的腿的数目的集合. 我可以想到有 0、2、4、6 和 8 条腿的动物, 以及一些有更多条腿的小动物. 如果你还想到了个别的像失去一条或多条腿的动物, 那你也可以将 1、3、5 和 7 等其他可能的数加入其值域. 不管怎样, 这个函数的值域并不是很清晰. 要想了解真实的答案, 你或许得是一位生物学家.
- 最后, 如果 $j(x)$ 是 Junkster 吃 x 时呕吐物的颜色, 那么其值域就会包含所有可能的呕吐物的颜色. 我很怕去想它们会是什么样的, 但或许亮蓝色不在其中吧.

1.1.1 区间表示法

在本书剩余部分, 函数总有上域 $\mathbb{R}$, 并且其定义域总会尽可能和 $\mathbb{R}$ 差不多 (除非另有说明). 因此, 我们会经常涉及实轴的子集, 尤其是像 $\{x : 2 \leqslant x < 5\}$ 这样的连通区间. 像这样写出完整的集合有点儿烦, 但总比说 "介于 2 和 5 之间的所有数, 包括 2 但不包括 5" 要强. 使用区间表示法会让我们做得更好.

我们约定, $[a,\ b]$ 是指从 a 到 b 端点间的所有实数, 包括 a 和 b. 所以 $[a,\ b]$ 指的是所有使得 $a \leqslant x \leqslant b$ 成立的 x 的集合. 例如, $[2,\ 5]$ 是所有介于 2 和 5 之间 (包括 2 和 5) 的实数的集合. (它不仅仅包括 2、3、4 和 5, 不要忘记还有一大堆处于 2 和 5 之间的分数和无理数, 比如 5/2、$\sqrt{7}$ 和 π.) 像 $[a,\ b]$ 这种形式表示的区间我们称作**闭区间**.

如果你不想包括端点, 把方括号变为圆括号就行了. 所以 $(a,\ b)$ 指的是介于 a 和 b 之间但不包括 a 和 b 的所有实数的集合. 这样, 如果 x 在区间 $(a,\ b)$ 中, 我们就知道 $a < x < b$. 集合 $(2,\ 5)$ 表示介于 2 和 5 之间但不包括 2 和 5 的所有实数的集合. 像 $(a,\ b)$ 这种形式表示的区间称作**开区间**.

你也可以混和匹配: $[a,\ b)$ 指的是介于 a 和 b 之间、包括 a 但不包括 b 的所有实数的集合; $(a,\ b]$ 包括 b, 但不包括 a. 这些区间在一个端点处是闭的, 而在另一个端点处是开的. 有时候, 像这样的区间称作**半开区间**. 上述的 $\{x : 2 \leqslant x < 5\}$ 就是一个例子, 也可以写成 $[2, 5)$.

还有一个有用的记号就是 (a, ∞), 它是指大于 a 但不包括 a 的所有数; $[a, \infty)$ 也一样, 只是它包括 a. 此外还有三个涉及 $-\infty$ 的可能性. 总而言之,各种情况如下.

$(a,\ b)$	$\{x: a < x < b\}$	a, b
$[a,\ b]$	$\{x: a \leqslant x \leqslant b\}$	a, b
$(a,\ b]$	$\{x: a < x \leqslant b\}$	a, b
$[a,\ b)$	$\{x: a \leqslant x < b\}$	a, b
$(a,\ \infty)$	$\{x: x > a\}$	a
$[a,\ \infty)$	$\{x: x \geqslant a\}$	a
$(-\infty,\ b)$	$\{x: x < b\}$	b
$(-\infty,\ b]$	$\{x: x \leqslant b\}$	b
$(-\infty,\ \infty)$	$\mathbb{R}$	

1.1.2 求定义域

有时候, 函数的定义中包括了定义域. (例如, 1.1 节中的函数 g 就是如此.) 然而在大多数情况下, 定义域是没有给出的. 通常的惯例是, 定义域包括实数集尽可能多的部分. 例如 $k(x) = \sqrt{x}$, 其定义域就不可能是 $\mathbb{R}$ 中的所有实数, 因为不可能

得到一个负数的平方根. 其定义域一定是 $[0, \infty)$, 就是大于或等于 0 的所有实数的集合.

好了, 我们知道取负数的平方根会出问题. 那么还有什么会把问题搞糟呢? 以下是三种最常见的情况.

(1) 分数的分母不能是零.

(2) 不能取一个负数的平方根 (或四次根, 六次根, 等等).

(3) 不能取一个负数或零的对数. (还记得对数函数吗? 若忘了, 请看看第 9 章!)

或许你还记得 $\tan(90°)$ 也是一个问题, 但这实际上是上述第一种情况的特例. 你看,

$$\tan(90°) = \frac{\sin(90°)}{\cos(90°)} = \frac{1}{0},$$

$\tan(90°)$ 之所以是无定义的, 实际上是因为其隐藏的分母为零. 这里还有一个例子: 如果定义

$$f(x) = \frac{\log_{10}(x+8)\sqrt{26-2x}}{(x-2)(x+19)},$$

那么 f 的定义域是什么呢? 当然, 为了使 $f(x)$ 有意义, 以下是我们必须要做的.

- 取 $(26-2x)$ 的平方根, 所以这个量必须是非负的. 也就是说, $26-2x \geqslant 0$. 这可以写成 $x \leqslant 13$.
- 取 $(x+8)$ 的对数, 所以这个量必须是正的. (注意对数和平方根的区别: 可以取 0 的平方根, 但不能取 0 的对数.) 不管怎么说, 我们需要 $x+8>0$, 所以 $x>-8$. 到现在为止, 我们知道 $-8<x\leqslant 13$, 所以其定义域最多是 $(-8, 13]$.
- 分母不能为 0, 这就是说 $(x-2)\neq 0$ 且 $(x+19)\neq 0$. 换句话说, $x\neq 2$ 且 $x\neq -19$. 最后一个条件不是问题, 因为我们已经知道 x 处于 $(-8,13]$ 内, 所以 x 不可能是 -19. 不过, 我们确实应该把 2 去掉.

这样就找到了其定义域是除了 2 以外的集合 $(-8, 13]$. 这个集合可以写作 $(-8, 13]\setminus\{2\}$, 这里的反斜杠表示 “不包括”.

1.1.3　利用图像求值域

让我们来定义一个新的函数 F, 指定其定义域为 $[-2, 1]$, 并且 $F(x)=x^2$ 在此定义域上. (记住, 我们看到的任何函数的上域总是所有实数的集合.) 同时又是对于所有的实数 x, $f(x)=x^2$. 那么 F 和 f 是同一个函数吗? 回答是否定的, 因为两个函数的定义域不相同 (尽管它们有相同的函数规则). 正如 1.1 节中的函数 g, 函数 F 是由限制 f 的定义域得到的.

现在, F 的值域又是什么呢? 如果你将 -2 到 1 之间 (包括 -2 和 1) 的每一个实数平方的话, 会发生什么呢? 你应该有能力直接求解, 但这是观察如何利用图像

来求一个函数的值域的很好机会. 基本思想是, 画出函数图像, 然后想象从图像的左边和右边很远的地方朝向 y 轴水平地射入两束亮光. 曲线会在 y 轴上有两个影子, 一个在 y 轴的左侧, 另一个在 y 轴的右侧. 值域就是影子的并集; 也就是说, 如果 y 轴上的任意一点落在左侧或右侧的影子里, 那么它处于函数的值域中. 我们以函数 F 为例来看一下这是怎么运作的吧.

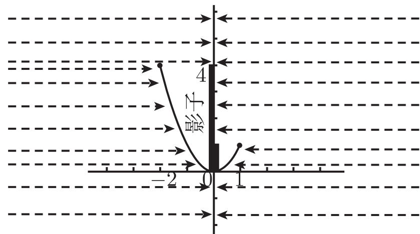


图　1-1

图 1-1 中左侧的影子覆盖了 y 轴从 0 到 4 (包括 0 和 4) 的所有点, 也就是 $[0, 4]$; 另一方面, 右侧的影子覆盖了从 0 到 1 (包括 0 和 1)的所有点, 也就是 $[0, 1]$. 右侧的影子没有贡献更多, 全部的覆盖范围仍然是 $[0, 4]$. 这就是函数F的值域.

1.1.4 垂线检验

在上一节中, 我们利用一个函数的图像来求其值域. 函数的图像非常重要：它真正地展示了函数 "看起来是什么样子的". 在第 12 章, 我们将会看到绘制函数图像的各种技巧, 但现在, 我很想提醒你注意的是垂线检验.

你可以在坐标平面上画任何你想画的图形, 但结果可能不是一个函数的图像. 那么函数的图像有什么特别之处呢？或者说, 什么是函数 f 的图像呢？它是所有坐标为 $(x, f(x))$ 的点的集合, 其中 x 在 f 的定义域中. 还有另外一种方式来看待它. 我们以某个实数 x 开始. 如果 x 在定义域中, 你就画点 $(x, f(x))$, 当然这个点在 x 轴上的点 x 的正上方, 高度为 $f(x)$. 如果 x 没有在定义域中, 你不能画任何点. 现在, 对于每一个实数 x, 我们重复这个过程, 从而构造出函数的图像.

这里的关键思想是, 你不可能有两个点有相同的 x 坐标. 换句话说, 在图像上没有两个点会落在相对于 x 轴的同一条垂线上. 要不然, 你又将如何知道在点 x 上方的两个或多个不同高度的点中, 哪一个是对应于 $f(x)$ 的值呢？这样就有了**垂线检验**：如果你有某个图像并想知道它是否是函数的图像, 你就看看是否任何的垂线和图像相交多于一次. 如果是这样的话, 那它就不是函数的图像; 反之, 如果没有一条垂线和图像相交多于一次, 那么你的确面对的是函数的图像. 例如, 以原点为中心, 半径为三个单位的圆的图像, 如图 1-2 所示.



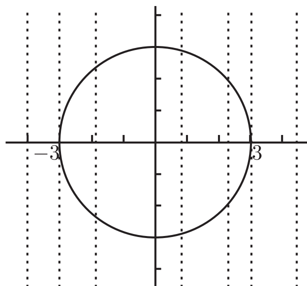


图　1-2

这么普通的对象应该是个函数, 对吗? 不对, 让我们进行如图所示的垂线检验. 当然, 在 -3 的左边或 3 的右边都没有问题 (垂线甚至都没有击中图像), 这很好. 就连在 -3 或 3 上, 垂线和图像也仅仅有一次相交, 这也很好. 问题出在 x 落在区间 $(-3,\ 3)$ 上时. 对于这其中的任意 x 值, 垂线通过 $(x,\ 0)$ 和圆相交两次, 这就坏事了. 你不知道 $f(x)$ 到底是对应上方的点还是下方的点.

最好的解决方法是把圆分成上下两个半圆, 并只选择上一半或者下一半. 整个圆的方程是 $x^2+y^2=9$, 而上半圆的方程是 $y=\sqrt{9-x^2}$, 下半圆的方程是 $y=-\sqrt{9-x^2}$. 这最后两个就是函数了, 定义域都是 $[-3,\ 3]$. 你可以以不同的方式来分割. 实际上, 你不是必须要把它分成半圆 (可以分割并改变上半圆和下半圆, 只要不违反垂线检验就行了). 例如, 图 1-3 也是一个函数的图像, 其定义域也是 $[-3,\ 3]$.

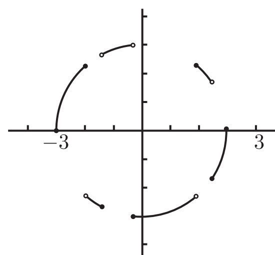


图　1-3

垂线检验通过, 所以这确实是一个函数的图像.

1.2　反　函　数

我们假设一个函数 f, 你给了它一个输入 x. 如果 x 在 f 的定义域中, 你就能得到一个输出, 我们称它为 $f(x)$. 现在, 我们把过程倒过来, 并问: 如果你选一个实数 y, 那么应该赋予 f 什么样的输入才能得到这个输出 y 呢?

用数学语言来陈述这个问题就是: 给定一个实数 y, 那么在 f 定义域中的哪个 x 满足 $f(x)=y$? 首先要注意的是, y 必须在 f 的值域中. 否则, 根据定义, 将不再有 x 的值使得 $f(x)=y$ 成立了. 如此在 f 定义域中将没有这样的 x 满足 $f(x)=y$, 因为值域是**所有**的可能输出.

另一方面, 如果 y 在值域当中, 也可能会有很多值都满足 $f(x)=y$. 例如 $f(x)=x^2$ (其定义域为 $\mathbb{R}$), 我们的问题是 x 取何值时会输出 64. 很显然, 有两个 x 值: 8 和 -8. 另外, 如果 $g(x)=x^3$, 对于相同的问题, 这时只有一个 x 值, 就是 4. 对于任意一个我们赋予 g 去做变换的实数, 结果都是如此, 因为任何数都只有一个 (实数) 立方根.

所以这里的情形如下: 给定一个函数 f, 在 f 的值域中选择 y. 在理想状况下, 仅有一个 x 值满足 $f(x)=y$. 如果上述理想状况对于值域中的每一个 y 来说都成

立, 那么就可以定义一个新的函数, 它将逆转变换. 从输出 y 出发, 这个新的函数发现一个且仅有一个输入 x 满足 $f(x)=y$. 这个新的函数称为 f 的**反函数**, 并写作 f^{-1}. 以下是使用数学语言对上述情形的总结.

(1) 从一个函数 f 出发, 使得对于在 f 值域中的任意 y, 都只有唯一的 x 值满足 $f(x)=y$. 也就是说, 不同的输入对应不同的输出. 现在, 我们就来定义反函数 f^{-1}.

(2) f^{-1} 的定义域和 f 的值域相同.

(3) f^{-1} 的值域和 f 的定义域相同.

(4) $f^{-1}(y)$ 的值就是满足 $f(x)=y$ 的 x. 所以,

$$\text{如果 } f(x)=y, \text{ 那么 } f^{-1}(y)=x.$$

变换 f^{-1} 就像是 f 的撤销按钮：如果你从 x 出发, 并通过函数 f 将它变换为 y, 那么你可以通过在 y 上的反函数 f^{-1} 来撤销这个变换的效果, 取回 x.

这会引发一些问题：你如何知道只有唯一的 x 值满足 $f(x)=y$ 呢？如果是这样, 如何求得反函数呢, 其图像又是什么样子呢？如果不是这样, 你又如何挽救这一局面呢？在接下来的三个小节中我们会对这些问题做出回答.

1.2.1 水平线检验

对于第一个问题 —— 如何知道对于 f 值域中的任意 y, 只有一个 x 值满足 $f(x)=y$ —— 最好的方法也许是看一下函数图像. 我们想要在 f 值域中选择 y, 并且希望只有一个 x 值满足 $f(x)=y$. 这就意味着通过点 $(0,y)$ 的水平线应该和图像仅有一次相交, 且交点为点 (x,y). 那个 x 就是我们想要的. 如果水平线和曲线相交多于一次, 那将会有多个可能的对应 x 值, 情况会很糟. 如果是那样, 获得反函数唯一的方法就是对定义域加以限制, 我们很快会讨论这一点. 如果水平线根本就没有和曲线相交, 会怎样呢？就是 y 根本没有在值域当中, 这样也不错.

这样一来, 就可以描述**水平线检验**：如果每一条水平线和一个函数的图像相交至多一次, 那么这个函数就有一个反函数. 如果即使只有一条水平线和图像相交多于一次, 那么这个函数就没有反函数. 例如, 我们来看一下图 1-4 中 $f(x)=x^3$ 和 $g(x)=x^2$ 的图像.

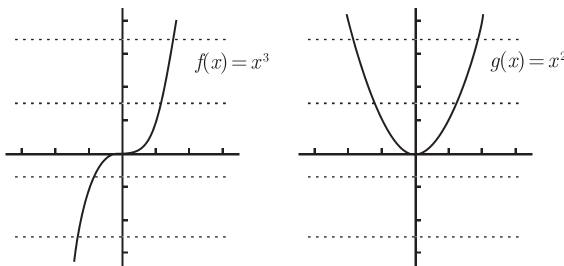


图 1-4

没有一条水平线和 $y = f(x)$ 相交多于一次, 所以 f 有一个反函数. 另一方面, 一些水平线和曲线 $y = g(x)$ 相交两次, 所以 g 没有反函数. 这里的问题在于：如果通过 $y = x^2$ 来求解 x, 其中 y 为正, 那么就会出现两个解：$x = \sqrt{y}$ 和 $x = -\sqrt{y}$. 结果你不知道该取哪一个.

1.2.2　求反函数

现在来看第二个问题：如何求得函数 f 的反函数呢？其实只需写下 $y = f(x)$, 然后试着解出 x. 在 $f(x) = x^3$ 的例子中, 有 $y = x^3$, 所以 $x = \sqrt[3]{y}$. 这就意味着, $f^{-1}(y) = \sqrt[3]{y}$. 如果你觉得变量 y 刺眼, 可以将它改写为 x, 写成 $f^{-1}(x) = \sqrt[3]{x}$.

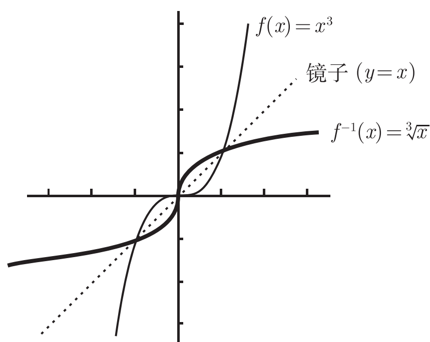


图　1-5

当然了, 求解 x 并不总是那么简单. 事实上, 求解经常是不可能的. 另一方面, 如果你知道函数图像是什么样子的, 反函数的图像就会很容易画出来. 基本思想是, 在图像上画一条 $y = x$ 的直线, 然后将这条直线假想为一个双面的镜子. 反函数就是原始函数的镜面反射. 如果 $f(x) = x^3$, 那么 f^{-1} 的图像如图 1-5 所示.

原始函数 f 在 $y = x$ 这面“镜子”中被反射, 从而得到反函数. 注意：f 和 f^{-1} 的定义域和值域都是整个实轴.

1.2.3　限制定义域

最后要处理第三个问题：如果水平线检验失败因而没有反函数, 那应该怎么办呢？我们面临的问题是, 对于相同的 y 有多个 x 值. 解决此问题的唯一方法是：除了这多个 x 值中的一个, 我们放弃所有其他值. 也就是说, 必须决定要保留哪一个 x 值, 然后放弃剩余的值. 正如我们在 1.1 节中看到的, 这称为**限制函数的定义域**. 实质上, 我们删去部分曲线, 使得保留下来的部分能够通过水平线检验. 例如 $g(x) = x^2$, 可以删除左半边的图像, 如图 1-6 所示.

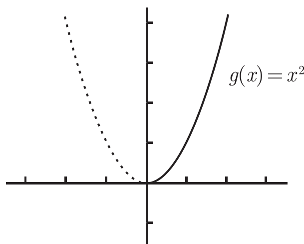


图　1-6

这条新的 (实线的) 曲线将定义域缩减为 $[0, \infty)$, 并且满足水平线检验, 所以它有反函数. 更确切地说, 定义在定义域 $[0, \infty)$ 上的函数 h 有反函数, 其中 $h(x) = x^2$. 让我们用镜面反射游戏来看一下它到底是什么样子的, 如图 1-7 所示.

为了找到反函数的方程, 我们必须在方程 $y = x^2$ 中解出 x. 很明显, 问题

的解就是 $x = \sqrt{y}$ 或 $x = -\sqrt{y}$, 但是我们需要哪一个呢？我们知道反函数的值域和原始函数的定义域是相同的, 而后者被限制为 $[0,\infty)$, 所以我们需要一个非负的数来作为答案, 即 $x = \sqrt{y}$. 这就是说, $h^{-1}(y) = \sqrt{y}$. 当然, 也可以把原始图像的右半边删除, 将定义域限制为 $(-\infty,\ 0]$. 在那种情况下, 我们得到一个定义域为 $(-\infty,\ 0]$ 的函数 j. 它也满足 $j(x) = x^2$, 但只是在这个定义域上才成立. 这个函数也有反函数, 反函数是负的平方根, 即 $j^{-1}(y) = -\sqrt{y}$.

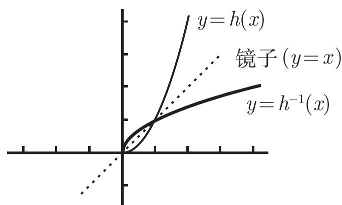
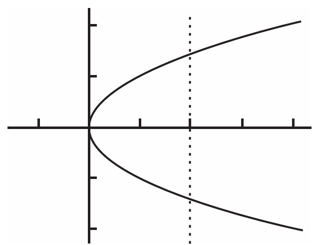


图 1-7


图 1-8

顺便说一下, 如果你让没有通过水平线检验的、定义域为 $(-\infty,\ \infty)$ 的原始函数 $g(x) = x^2$ 在镜子 $y = x$ 中反射, 那么你会得到如图 1-8 所示的图像.

注意到这个图像不会通过垂线检验, 所以它不是函数的图像. 这说明了垂线检验和水平线检验之间的联系, 即水平线被镜子 $y = x$ 反射后会变成垂线.

1.2.4 反函数的反函数

有关反函数还有一点：如果 f 有反函数, 那么对于在 f 定义域中的所有 x, $f^{-1}(f(x)) = x$ 成立; 同样, 对于在 f 值域当中的所有 y, 都有 $f\left(f^{-1}(y)\right) = y$. (记得, f 的值域和 f^{-1} 的定义域相同, 所以对于 f 值域中的 y, 我们确实可以取到 $f^{-1}(y)$, 不会导致任何曲解.)



例如 $f(x) = x^3$, f 的反函数由 $f^{-1}(x) = \sqrt[3]{x}$ 给出, 所以对于任意的 x, $f^{-1}(f(x)) = \sqrt[3]{x^3} = x$. 不要忘记, 反函数就像是撤销按钮. 我们使用 x 作为 f 的输入, 然后给出输出到 f^{-1}; 这撤销了变换并让我们取回了 x 这个原始的数. 类似地, $f\left(f^{-1}(y)\right) = \left(\sqrt[3]{y}\right)^3$. 所以, f^{-1} 是 f 的反函数, 且 f 是 f^{-1} 的反函数. 换句话说, 反函数的反函数就是原始函数.

不过, 对于限制定义域的情况一定要当心. 令 $g(x) = x^2$, 我们已经看到你需要对其定义域加以限制, 方能取得反函数. 设想我们把定义域限制为 $[0,\ \infty)$, 但由于粗心大意而把函数继续看成是 g 而不是先前小节中那样的 h. 我们便会说 $g^{-1}(x) = \sqrt{x}$. 如果你真要计算 $g\left(g^{-1}(x)\right)$, 你就会发现它是 $(\sqrt{x})^2$, 即等于 x, 只要 $x \geqslant 0$. (当然, 不是这样的话, 从一开始你就无法取得平方根.)

另一方面, 如果你解出 $g^{-1}(g(x))$, 你会得到 $\sqrt{x^2}$, 它不是总和 x 相同. 例如, 如果 $x = -2$, 那么 $x^2 = 4$, $\sqrt{x^2} = \sqrt{4} = 2$. 所以一般而言, $g^{-1}(g(x)) = x$ 不成立.

这里的问题在于, -2 没有在 g 的限制定义域当中. 而且, 从技术角度而言, 你甚至不可能计算 $g(-2)$, 因为 -2 不再属于 g 的定义域了. 我们确实应该使用 h, 而不是 g, 以便提醒自己要更加小心. 不过在实践中, 数学家们在限制定义域时经常不会改变字母! 所以把这种情形总结如下对大家是很有帮助的.



如果一个函数 f 的定义域可以被限制, 使得 f 有反函数 f^{-1}, 那么

- 对于 f 值域中的所有 y, 都有 $f\left(f^{-1}(y)\right) = y$; 但是
- $f^{-1}(f(x))$ 可能不等于 x; 事实上, $f^{-1}(f(x)) = x$ 仅当 x 在限制的定义域中才成立.

在 10.2.6 节, 对于反三角函数, 我们会再次提到这些要点.

1.3　函数的复合

假设有一个表达式为 $g(x) = x^2$ 的函数 g. 你可以将 x 替换成任何使函数有意义的对象, 如 $g(y) = y^2$ 或 $g(x+5) = (x+5)^2$. 后一个例子需要特别注意小括号, 若写成 $g(x+5) = x+5^2$ 就是错的, 因为 $x+25$ 并不等于 $(x+5)^2$. 所以在替换过程中如果拿不准, 可用小括号. 也就是说, 如果你需将 $f(x)$ 写成 f(某表达式), 可将每一个 x 替换成 (某表达式), 这时一定要加小括号. 唯一不需要加小括号的情况是, 当函数是指数函数时, 如 $h(x) = 3^x$, 你可以写成 $h(x^2+6) = 3^{x^2+6}$. 不需要加小括号是因为你已经将 x^2+6 写成上标了.

现在考虑定义为 $f(x) = \cos(x^2)$ 的函数 f. 若给定一个数 x, 如何计算 $f(x)$ 呢? 你会首先计算 x 的平方, 然后计算平方值的余弦. 鉴于我们可将 $f(x)$ 的计算分解成前后相继的两个独立的计算, 我们也就可以将这些计算各描述成一个函数. 因此, 令 $g(x) = x^2$, $h(x) = \cos(x)$. 为了模拟函数 f 是如何作用于输入值 x 的, 你可先将 x 输入到函数 g 进行求平方运算, 接着不必返回 g 的结果而直接让 g 将其结果作为函数 h 的输入, 然后 h 计算出一个最终的结果值, 该结果值当然是由函数 g 计算出的 x 平方值的余弦值. 这个过程恰恰模拟了 f, 故我们可以写出 $f(x) = h(g(x))$, 也可表示为 $f = h \circ g$, 这里的圈表示 "与 …… 的复合", 即 f 是 g 与 h 的**复合**. 换言之, f 是 g 与 h 的复合函数. 这里需要小心的是, 我们把 h 写在 g 的前面 (像平常一样从左向右读), 但计算时我们要先从 g 开始. 我承认这确实容易让人搞混, 但我也没办法 —— 你只能试着去接受.



练习求两个或多个函数的复合是很有用的. 例如, 若 $g(x) = 2^x, h(x) = 5x^4$, $j(x) = 2x-1$, 则函数 $f = g \circ h \circ j$ 的表达式是什么? 我们只需从 j 开始, 将其代换到 h, 接着再将结果代换到 g, 可得

$$f(x) = g(h(j(x))) = g(h(2x-1)) = g(5(2x-1)^4) = 2^{5(2x-1)^4}.$$



同样, 你需要练习该过程的逆过程. 例如, 假定你开始于函数

$$f(x) = \frac{1}{\tan(5\log_2(x+3))}.$$

如何将 f 分解为几个简单函数呢？从函数式中找到 x, 首先需要加 3, 所以设 $g(x) = x+3$; 然后要对所得值取以 2 为底的对数, 所以令 $h(x) = \log_2(x)$; 接着需乘 5, 则设 $j(x) = 5x$; 再接着要求正切值, 因此令 $k(x) = \tan(x)$; 最后要取倒数, 于是令 $m(x) = 1/x$. 由上, 验证下式：

$$f(x) = m(k(j(h(g(x))))).$$

利用复合符号, 可以写成

$$f = m \circ k \circ j \circ h \circ g.$$

这并不是函数 f 的唯一分解形式. 例如, 我们可以将函数 h 和 j 复合成另一个函数 n, 其中 $n(x) = 5\log_2(x)$. 然后你应该验证一下 $n = j \circ h$ 和

$$f = m \circ k \circ n \circ g.$$

或许最初 (包含 j 和 h) 的分解较好一点, 因为它将 f 分解成更多的基本形式, 但第二种 (包含 n) 也没错, 毕竟 $n(x) = 5\log_2(x)$ 仍是关于 x 的较为简单的函数.

注意, 函数的复合并不是把它们相乘. 例如 $f(x) = x^2\sin(x)$, f 不是两个函数的复合, 因为对任意给定的 x, 计算 $f(x)$ 的值需要求解 x^2 和 $\sin(x)$(先求哪个值都没关系, 这与复合函数不同), 然后将这两个值乘起来. 若令 $g(x) = x^2$, $h(x) = \sin(x)$, 则我们可以写成 $f(x) = g(x)h(x)$ 或 $f = gh$. 可将它与这两个函数的复合函数 $j = g \circ h$, 即

$$j(x) = g(h(x)) = g(\sin(x)) = (\sin(x))^2$$

或 $j(x) = \sin^2(x)$ 比较一下. 函数 j 完全不同于乘积 $x^2\sin(x)$, 它同样不同于函数 $k = h \circ g$. 函数 k 也是 g 和 h 的复合函数, 不过是按另一个顺序的复合：

$$k(x) = h(g(x)) = h(x^2) = \sin(x^2).$$

k 是另一个完全不同的函数. 这个例子说明, 函数的乘积和复合是不同的, 且函数的复合与函数顺序有关系, 而函数的乘积与函数顺序无关.

复合函数另一个简单但重要的例子是, 将函数 f 和 $g(x) = x - a$(a 是常数) 进行复合. 对复合得到的新函数 $h(x) = f(x-a)$, 需要关注的是新函数 $y = h(x)$ 和函数 $y = f(x)$ 的图像是一样的, 只不过 $y = h(x)$ 的函数图像向右平移了 a 个单位. 如果 a 是负的, 那么就是向左平移. (一种理解方式是, 向右平移 -3 个单位与向左平移 3 个单位是一样的.) 那么如何画 $y = (x-1)^2$ 的图像呢？就像画 $y = x^2$ 的图像一样, 只是用 $x - 1$ 来代替 x. 所以可将函数 $y = x^2$ 的图像向右平移 1 个单位, 如图 1-9 所示.

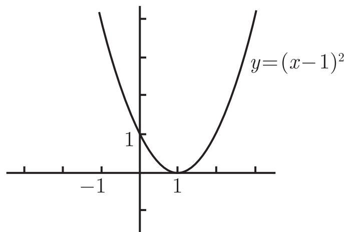


图 1-9



类似地, $y=(x+2)^2$ 的图像是将 $y=x^2$ 的图像向左平移 2 个单位, 可把 $(x+2)$ 理解为 $(x-(-2))$.

1.4 奇函数和偶函数

一些函数具有对称性, 这便于对它们进行讨论. 考虑定义为 $f(x)=x^2$ 的函数 f, 任选一个正数 (我选 3) 作用于函数 f(得到 9). 现在取该数的负值, 由我选择的数可得 -3, 将其作用于函数 f(又得到 9). 不论你选择的是几, 应该跟我一样, 两次得到了相同的值. 你可将这种现象表示为, 对所有的 x, 有 $f(-x)=f(x)$. 也就是说, 将 x 作为 f 的输入和将 $-x$ 作为输入, 会得到一样的结果. 注意到 $g(x)=x^4$ 和 $h(x)=x^6$ 同样具有这种性质. 事实上, 当 n 是偶数时 (n 可以是负数), $j(x)=x^n$ 具有相同的性质. 受以上讨论的启发, 我们说, 如果对 f 定义域里的所有 x 有 $f(-x)=f(x)$, 则 f 是偶函数. 这个等式对某些 x 值成立是不够的, 它必须对定义域里的**所有**x 都成立.

现在, 我们对函数 $f(x)=x^3$ 做相同的讨论. 选择你喜欢的任一正数 (我仍选 3) 作用于 f (得到 27). 用你选的数的负值再试一遍, 我的数的负值是 -3, 得到 -27, 你同样应该得到先前结果的负值. 可以用数学方式将其表示为 $f(-x)=-f(x)$. 同样地, 当 n 是奇数时 (n 可以是负数), $j(x)=x^n$ 具有相同的性质. 因此我们说, 当对 f 定义域内所有 x 都有 $f(-x)=-f(x)$ 时, f 是奇函数.

一般而言, 一个函数可能是奇的, 可能是偶的, 也可能非奇非偶. 要记住这一点, 大多数函数是非奇非偶的. 另一方面, 只有一个函数是既奇又偶的, 它就是非常单调的对所有 x 都成立的 $f(x)=0$(我们称之为零函数). 它为什么是唯一的既奇又偶的函数呢? 我们证明一下. 若函数 f 是偶函数, 则对所有 x 有 $f(-x)=f(x)$; 但如果同时它又是奇的, 则对所有 x 有 $f(-x)=-f(x)$, 用第一个等式减去第二个等式, 得到 $0=2f(x)$, 即 $f(x)=0$, 这对所有 x 成立, 因此函数 f 一定是零函数. 另一个有用的结论是, 如果一个函数是奇的, 并且 0 在其定义域内, 则 $f(0)=0$. 为什么呢? 由于对定义域里的所有 x, f 都有 $f(-x)=-f(x)$, 我们用 0 试一下. 我们得 $f(-0)=-f(0)$, 但 -0 等于 0, 因此 $f(0)=-f(0)$, 化简得 $2f(0)=0$, 即 $f(0)=0$.

不论如何, 对于一个函数 f, 怎么来判定它是奇函数、偶函数或都不是呢? 若是奇函数或偶函数又怎样呢? 我们先来看下第二个问题, 然后再讨论第一个问题. 当知道一个函数的奇偶性之后, 一个比较好的事情就是画函数图像比较容易了. 事实上, 如果你能将这个函数的右半边图像画出来, 那么画左半边图像就是小菜一碟. 我们先讨论当 f 是偶函数时的情形. 因 $f(x)=f(-x)$, $y=f(x)$ 的图像在 x 和 $-x$ 坐标上方具有相同的高度, 且对所有的 x 都成立, 如图 1-10 所示.

我们得到这样的结论：**偶函数的图像关于 y 轴具有镜面对称性**. 所以当你画

出偶函数的右半边图像后, 就可以通过将其图像关于 y 轴反射得到它的左半边图像. 不妨用 $y=x^2$ 的图像检验一下它的镜面对称性.

另一方面, 假设 f 是奇函数. 因 $f(-x)=-f(x)$, $y=f(x)$ 图像在 x 坐标**上方**和 $-x$ 坐标**下方**具有相同的高度. (当然, 若 $f(x)$ 是负的, 你可以调换一下 "上方" 和 "下方" 两个词.) 不论如何, 其图像如图 1-11 所示.

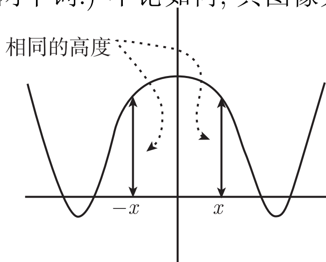


图 1-10

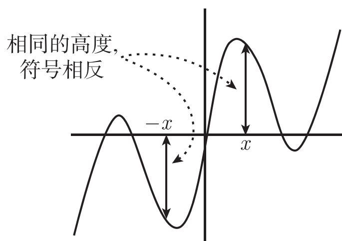


图 1-11

现在的对称性是关于原点的点对称, 即**奇函数的图像关于原点有 180° 的点对称性**. 这就意味着, 如果你只有奇函数的右半边图像, 你可按下面的方法得到其左半边的图像. 想象该曲线是浮在纸面上, 你能够把它拿起来但不能改变它的形状. 不过, 你没有把它拿起来, 而是用大头针在原点处把曲线钉住 (回想一下, 奇函数若在 0 处有定义, 它必定通过原点), 然后将整个曲线旋转半圈, 这样就得到左半边图像的样子了. (如果曲线是不连续的, 即不是连在一起的一条, 这个方法就不那么好用了.) 可验证一下, 上面的图像和函数 $y=x^3$ 的图像都具有这样的对称性.

现在假设 f 定义为 $f(x)=\log_5(2x^6-6x^2+3)$, 你怎么确定 f 是奇函数、偶函数, 还是都不是呢? 方法就是, 将每个 x 替换为 $(-x)$ 并计算 $f(-x)$, 一定要记着给 $-x$ 加上小括号, 然后化简结果. 如果你得出了原始表达式 $f(x)$, f 就是偶的; 如果得到原始表达式的负值 $-f(x)$, f 就是奇的; 如果得到的结果一团糟, 既不是 $f(x)$ 也不是 $-f(x)$, 则 f 就非奇非偶 (或之前的化简不充分). 由上例, 可得

$$f(-x)=\log_5(2(-x)^6-6(-x)^2+3)=\log_5(2x^6-6x^2+3),$$

本式实际上等于 $f(x)$ 本身, 因此函数 f 是偶的. 那函数

$$g(x)=\frac{2x^3+x}{3x^2+5} \quad \text{和} \quad h(x)=\frac{2x^3+x-1}{3x^2+5}$$

的奇偶性又如何呢? 对函数 g, 我们有

$$g(-x)=\frac{2(-x)^3+(-x)}{3(-x)^2+5}=\frac{-2x^3-x}{3x^2+5}.$$

现在可把负号提到前面来, 得到

$$g(-x)=-\frac{2x^3+x}{3x^2+5},$$

注意到结果等于 $-g(x)$, 即除了负号以外, 剩下部分就是原始函数, 因此 g 是奇函数. 那函数 h 呢? 我们有

$$h(-x)=\frac{2(-x)^3+(-x)-1}{3(-x)^2+5}=\frac{-2x^3-x-1}{3x^2+5}.$$

我们再次把负号提到前面来, 得到

$$h(-x)=-\frac{2x^3+x+1}{3x^2+5}.$$

嗯, 看起来这不是原始函数的负值, 因为分子上有个 $+1$. 它也不是原始函数本身, 所以函数 h 是非奇非偶的.

我们再看一个例子. 若想证明两个奇函数之积是偶函数, 该怎么做呢? 先给事物命名比较利于讨论, 我们就定义有两个奇函数 f 和 g. 我们需要看一下它们的乘积, 因此定义它们的积为 h, 即定义了 $h(x)=f(x)g(x)$, 而我们的任务是要证明 h 是偶的. 像往常一样, 我们需要证明 $h(-x)=h(x)$. 因 f 和 g 都是奇的, 注意到 $f(-x)=-f(x)$, $g(-x)=-g(x)$ 会有所帮助. 我们从 $h(-x)$ 开始. 由于 h 是 f 和 g 的乘积, 有 $h(-x)=f(-x)g(-x)$. 再利用 f 和 g 的奇函数性质将等式右边表示为 $(-f(x))(-g(x))$, 负号提到前面消掉, 由此得到 $f(x)g(x)$, 而它当然等于 $h(x)$. 我们可以 (也应该) 把上述过程用数学式表示为

$$h(-x)=f(-x)g(-x)=(-f(x))(-g(x))=f(x)g(x)=h(x).$$

总之, 由 $h(-x)=h(x)$ 可得函数 h 是偶函数. 现在你应该可以证明两偶函数之积仍为偶函数, 奇函数和偶函数之积是奇函数. 马上试一下吧!

1.5　线性函数的图像

形如 $f(x)=mx+b$ 的函数叫作**线性函数**. 如此命名原因很简单, 因为它们的图像是直线. 直线的斜率是 m. 设想一下, 此时此刻你就在这页纸中, 这条直线就像是座山, 你从左向右开始登山, 如图 1-12 所示.

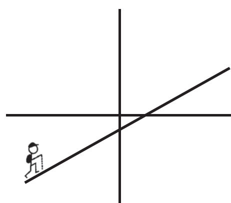


图　1-12

如果像左图一样, 斜率 m 为正数, 那么你正在上山. m 越大, 这段上坡就越陡. 相反, 如果 m 为负数, 那么你正在下山. m 的数值越小 (即绝对值越大), 这段下坡也就越陡. 如果斜率为 0, 这段山路就是水平的, 你既不在上山, 也不在下山, 仅仅是在沿一条水平直线前行.

你仅仅需要确认两个点, 就可以画出线性函数的图像, 因为两点确定一条直线. 你所要做的就是把尺子放在这两点上, 笔轻轻一连就行了. 其中一点很容易找, 就是 y 轴的截距. 设 $x=0$, 很显然 $y=m\times 0+b=b$. 也就是说, y 轴的截距为 b, 所以直线通过 $(0,b)$ 这点. 我们可以通过找 x 轴的截距来找另一点, 设 y 为 0, 求 x

的值. 不过, 这种方法在两种特殊情况下不适用.

情况一：$b = 0$, 这时函数变为 $y = mx$. 直线通过原点, x 轴和 y 轴的截距都为零. 为了求得另一点, 可以把 $x = 1$ 代入, 可得 $y = m$. 所以直线 $y = mx$ 通过原点和 $(1,\ m)$ 这两点. 例如, 直线 $y = -2x$ 通过原点和 $(1, -2)$, 如图 1-13 所示.

情况二：当 $m = 0$, 这时函数变为 $y = b$, 是一条通过 $(0, b)$ 的水平直线.

更有趣的例子, 可考虑函数 $y = \frac{1}{2}x - 1$. 很显然, y 轴截距为 -1, 斜率为1/2. 为画这条直线, 我们还需要求出 x 轴的截距. 通过设 $y = 0$ 可以得出 $0 = \frac{1}{2}x - 1$, 化简后得出 $x = 2$. 图像如图 1-14 所示.

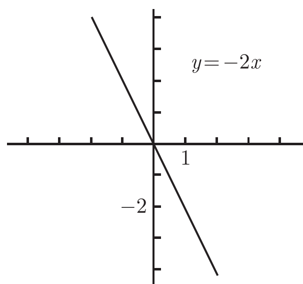


图 1-13

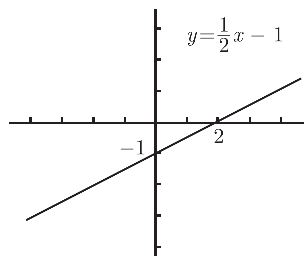


图 1-14

现在假设你知道平面上有一条直线, 但不知道它的方程. 如果你知道这条直线通过某一固定的点以及它的斜率, 那就能很容易地找到它的方程. 你真的, 真的, 真的, 很有必要去掌握这种方法, 因为它经常出现. 这个公式叫直线方程的**点斜式**, 其文字表达如下：

> 如果已知直线通过点 $(x_0,\ y_0)$, 斜率为 m, 则它的方程为 $y - y_0 = m(x - x_0)$.

如果已知一条直线通过 $(-2, 5)$, 斜率为 -3, 如何求它的方程？方程为 $y - 5 = -3(x - (-2))$, 化简后结果为 $y = -3x - 1$.

有时你不知道直线的斜率, 但知道它通过哪两点. 那怎样求它的方程呢？技巧是, 找出它的斜率, 再用刚才的方法去求出方程. 首先, 你需要知道：

> 如果一条直线通过点 $(x_1,\ y_1)$ 和 $(x_2,\ y_2)$, 则它的斜率等于 $\frac{y_2 - y_1}{x_2 - x_1}$.

例如, 通过 $(-3, 4)$ 和 $(2, -6)$ 的直线方程是什么？首先, 求它的斜率：

$$斜率 = \frac{-6 - 4}{2 - (-3)} = \frac{-10}{5} = -2.$$

我们现在知道该直线通过 $(-3, 4)$, 斜率为 -2, 所以它的方程为 $y - 4 = -2(x - (-3))$, 化简后为 $y = -2x - 2$. 同样, 我们也可以使用另一点 $(2, -6)$ 和斜率为 -2, 得出方

程为 $y-(-6)=-2(x-2)$, 化简后为 $y=-2x-2$. 你会发现, 无论使用哪一个点, 最后得到的结果都是相同的.

1.6 常见函数及其图像

下面是你应该知道的最重要的一些函数.

(1) **多项式** 有许多函数是基于 x 的非负次幂建立起来的. 你可以以 1、x、x^2、x^3 等为基本项, 然后用实数同这些基本项做乘法, 最后把有限个这样的项加到一起. 例如, 多项式 $f(x)=5x^4-4x^3+10$ 是由 x^4 的 5 倍加 x^3 的 -4 倍加 10 而形成的. 你可能也想加中间的基本项 x^2 和 x , 但由于它们没有出现, 所以我们可以说零倍的 x^2 和零倍的 x. 基本项 x^n 的倍数叫作 **x^n 的系数**. 例如, 刚才的多项式 x^4、x^3、x^2、x 和常数项的系数分别为 5、-4、0、0 和 10. (顺便提一下, 为什么会有 x 和 1 的形式? 这两项看上去与其他项不同, 但它们实际上是一样的, 因为 $x=x^1, 1=x^0$.) 最大的幂指数 n (该项系数不能为零) 叫作多项式的**次数**. 例如上述多项式的次数为 4, 因为不存在比 4 大的 x 的幂指数. 次数为 n 的多项式的数学通式为

$$p(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_2x^2+a_1x+a_0,$$

其中 a_n 为 x^n 的系数, a_{n-1} 为 x^{n-1} 的系数, 以此类推, 直到最后一项 1 的系数为 a_0.

由于 x^n 是所有多项式的基本项, 因而你应该知道它们的图像是什么样的. 偶次幂的图像之间是非常类似的; 同样, 奇次幂的图像之间也很类似. 图 1-15 是从 x^0 到 x^7 的图像.

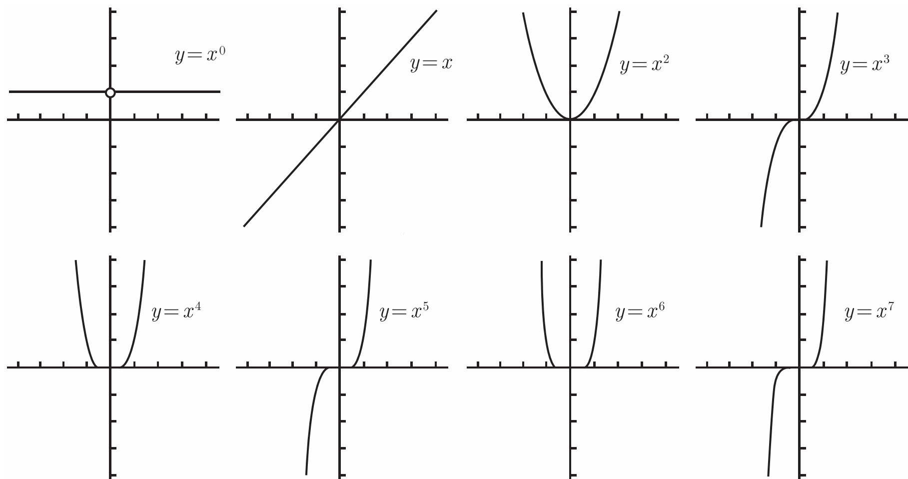
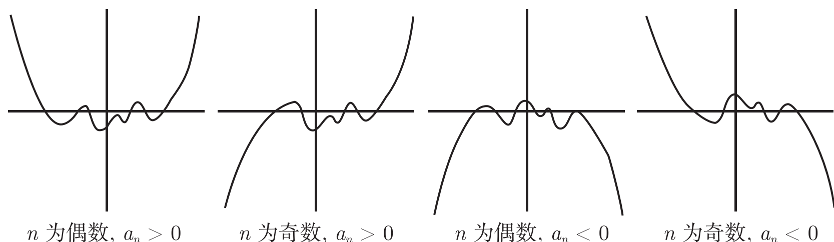


图 1-15

一般的多项式的图像是很难画的. 除非是很简单的多项式, 否则连 x 轴的截距

都经常很难找到. 不过, 多项式的图像左右两端的走势倒是容易判断. 这是由最高次数的项的系数决定的, 该系数叫作**首项系数**. a_n 就为上述多项式通式的首项系数. 例如, 我们刚才提到的多项式 $5x^4-4x^3+10$, 5 为它的首项系数. 实际上, 我们只需考虑首项系数正负以及多项式次数的奇偶就能判断图像两端的走势了. 所以图像两端的走势共有如下四种情况, 如图 1-16 所示.



n 为偶数, $a_n>0$ n 为奇数, $a_n>0$ n 为偶数, $a_n<0$ n 为奇数, $a_n<0$

图 1-16

上述图像的中间部分是由多项式的其他项决定的. 上图仅仅是为了显示图像左右两端的走势. 在这个意义上, 多项式 $5x^4-4x^3+10$ 的图像同最左边的图像类似, 因为 $n=4$ 为偶数, $a_n=5$ 为正数.

我们再稍微讨论一下次数为 2 的多项式, 又叫**二次函数**. 不写成 $p(x)=a_2x^2+a_1x+a_0$, 而把系数分别写成 a、b、c 会更简单些, 即我们有 $p(x)=ax^2+bx+c$. 根据判别式的符号可以判断二次函数到底有两个、一个还是没有实数解. 通常我们用希腊字母 Δ 来表示判别式 $\Delta=b^2-4ac$. 它共有三种可能性. 如果 $\Delta>0$, 有两个不同的解; 如果 $\Delta=0$, 只有一个解, 也可以说有两个相同的解; 如果 $\Delta<0$, 在实数范围内无解. 对于前两种情况, 解为

$$\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

注意到该表达式根号下为判别式. 二次函数的一个重要技术是**配方**. 下面举例说明. 考虑二次函数 $2x^2-3x+10$. 第一步把二次项的系数提出来, 多项式变为了 $2\left(x^2-\frac{3}{2}x+5\right)$. 这时就得到一个二次项系数为 1 的多项式. 接下来的关键一步是把 x 的系数, 这里是 $-\frac{3}{2}$, 除以 2, 再平方. 我们得到 $\frac{9}{16}$. 我们多希望常数项是 $\frac{9}{16}$, 而不是 5, 所以我们开动脑筋:

$$x^2-\frac{3}{2}x+5=x^2-\frac{3}{2}x+\frac{9}{16}+5-\frac{9}{16}.$$

为什么要加一次 $\frac{9}{16}$, 又减一次 $\frac{9}{16}$ 呢? 因为这样的话, 前三项为平方形式 $\left(x-\frac{3}{4}\right)^2$. 这时我们得到

$$x^2-\frac{3}{2}x+5=\left(x^2-\frac{3}{2}x+\frac{9}{16}\right)+5-\frac{9}{16}=\left(x-\frac{3}{4}\right)^2+5-\frac{9}{16}.$$

接下来, 只剩最后一小步, $5-\frac{9}{16}=\frac{71}{16}$. 最后恢复系数 2, 我们有

$$2x^2-3x+10=2\left(x^2-\frac{3}{2}x+5\right)=2\left(\left(x-\frac{3}{4}\right)^2+\frac{71}{16}\right)=2\left(x-\frac{3}{4}\right)^2+\frac{71}{8}.$$

事实证明, 这个形式在许多情形中更为便利. 你一定要学会如何配方, 因为我们要在第 18 章和第 19 章大量运用这个技巧.

(2) **有理函数**　形如 $\frac{p(x)}{q(x)}$, 其中 p 和 q 为多项式的函数, 叫作有理函数. 有理函数变化多样, 它的图像根据 p 和 q 两个多项式的变化而变化. 最简单的有理函数是多项式本身, 即 $q(x)$ 为 1 的有理函数. 另一个简单的例子是 $1/x^n$, 其中 n 为正整数. 图 1-17 是一些有理函数的图像.

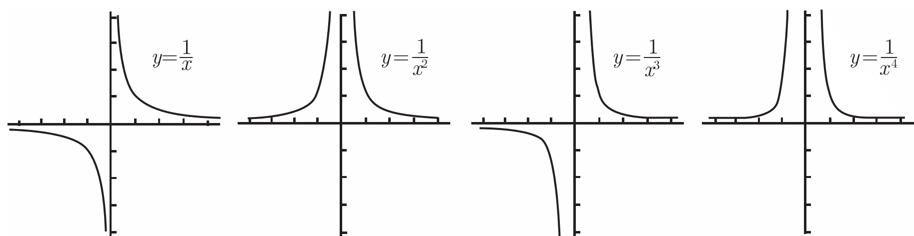


图　1-17

奇次幂的图像之间类似, 偶次幂的图像之间也很类似. 知道这些图像长什么样子是有帮助的.

(3) **指数函数和对数函数**　你需要知道指数函数的图像长什么样. 例如, 图 1-18 是 $y=2^x$ 的图像.

$y=b^x(b>1)$ 的图像都与这图类似. 有几点值得注意. 首先, 该函数的定义域为全体实数; 其次, y 轴的截距为 1 并且值域为大于零的实数; 最后, 左端的水平渐近线为 x 轴. 再强调一下, 该图像非常接近于 x 轴, 但永远不会接触到 x 轴, 无论在你的图形计算器上多么接近. (在第 3 章中, 我们会再次碰到渐近线.) $y=2^{-x}$ 的图像是 $y=2^x$ 关于 y 轴的对称, 如图 1-19 所示.

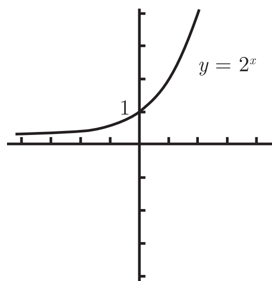


图　1-18

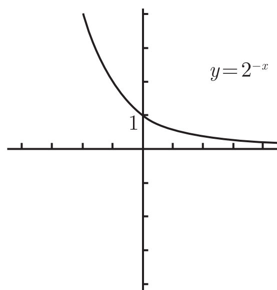


图　1-19

如果底小于 1, 情况会是怎样? 例如, 考虑 $y=\left(\frac{1}{2}\right)^x$ 的图像. 注意到 $\left(\frac{1}{2}\right)^x=1/2^x=2^{-x}$, 所以图 1-19 中 $y=2^{-x}$ 的图像也是 $y=\left(\frac{1}{2}\right)^x$ 的图像, 因为对于任意 $x, 2^{-x}$ 与 $\left(\frac{1}{2}\right)^x$ 均相等. 同理可得任何 $y=b^x(0<b<1)$ 的图像.

由于 $y=2^x$ 的图像满足水平线检验, 所以该函数有反函数. 这个反函数就是以 2 为底的对数函数 $y=\log_2(x)$. 以直线 $y=x$ 为镜子, $y=\log_2(x)$ 的图像如图 1-20 所示.

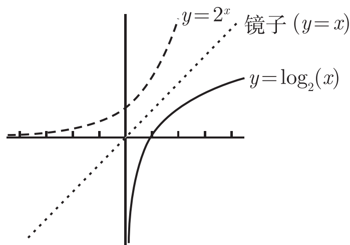


图 1-20

该函数的定义域为 $(0,+\infty)$, 这也印证了我之前所说的负数和 0 不能求对数的说法. 值域为全体实数, y 轴为垂直渐近线. $\log_b(x)(b>1)$ 的图像都很相似. 对数函数在微积分的学习中很重要, 你一定要学会怎样画它们的图像. 我们将在第 9 章学习对数函数的性质.

(4) **三角函数** 三角函数很重要, 所以下一章整章将对其作详细介绍.

(5) **带有绝对值的函数** 让我们看一下形如 $f(x)=|x|$ 的**绝对值函数**. 该函数的定义为:

$$|x|=\begin{cases} x & \text{如果 } x\geqslant 0,\\ -x & \text{如果 } x<0.\end{cases}$$

另一个看待这个绝对值函数的方法是, 它表示数轴上 0 和 x 的距离. 更一般而言, 你应该知道如下重要事实:

$|x-y|$ 是数轴上 x 和 y 两点间的距离.

例如, 假设你需要在数轴上找出区域 $|x-1|\leqslant 3$. 我们可以将该不等式阐释为 x 和 1 之间的距离小于或等于 3. 也就是说, 我们要找到所有与 1 之间的距离不大于 3 的点. 所以我们画一个数轴并标记 1 的位置, 如图 1-21 所示.





图 1-21

距离不大于 3 的点最左到 -2 最右到 4, 所以区域如图 1-22 所示.

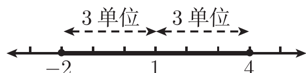


图 1-22

所以区域 $|x-1| \leqslant 3$ 也可表示为 $[-2, 4]$.

同样成立的是, $|x| = \sqrt{x^2}$. 可以检验一下. 当 $x \geqslant 0$, 显然 $\sqrt{x^2} = x$; 如果 $x < 0$, $\sqrt{x^2} = x$ 这个表达式就错了, 因为左边为正, 右边为负. 正确的表达式为 $\sqrt{x^2} = -x$, 这次右边为正了, 负负得正. 如果你再重新看一次 $|x|$ 的定义, 就会发现我们已经证明了 $|x| = \sqrt{x^2}$. 但尽管这样, 对于 $|x|$ 这个函数, 最好还是用分段函数去定义.

最后, 我们来看一些图像. 如果你知道一个函数的图像, 那么可以这样得到这个函数的绝对值的图像, 即以 x 轴为镜子, 把 x 轴下方的图像映射上来, x 轴上方的图像保持不变. 例如, 对于 $|x|$ 的图像, 可以通过翻转 $y = x$ 在 x 轴下方的部分得到, 图 $y = |x|$ 的图像如图 1-23 所示.

怎样画 $y = |\log_2(x)|$ 的图像呢? 使用图像对称的原理, 这个绝对值函数的图像如图 1-24 所示.

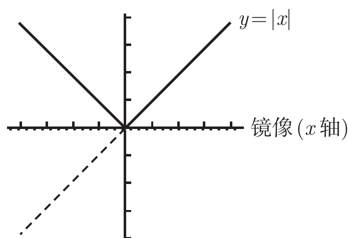


图　1-23

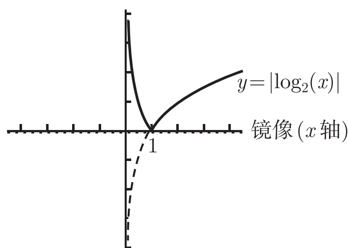


图　1-24

除了三角函数要在下一章讲外, 这是我在函数部分要讲解的所有内容. 但愿你之前已经见过本章中的许多内容, 因为其中的大部分知识将在微积分中被反复使用, 所以你需要尽快掌握这些知识.

第 2 章　三角学回顾

学习微积分必须要了解三角学. 说实话, 我们一开始不会碰到很多有关三角学的内容, 但当它们出现的时候, 会让我们感觉不容易. 因此, 我们不妨针对三角学最重要的一些方面进行一次全面的回顾:

- 用弧度度量的角与三角函数的基本知识;
- 实轴上的三角函数 (不只是介于 $0°$ 和 $90°$ 的角);
- 三角函数的图像;
- 三角恒等式.

准备开始回忆吧 ……

2.1　基本知识

首先要回忆的是弧度的概念. 旋转一周, 我们说成 2π **弧度**而不是 $360°$. 这似乎有点古怪, 但这里也有一个理由, 那就是半径为 1 个单位的圆的周长是 2π 个单位. 事实上, 这个圆的一个扇形的弧长就是这个扇形的圆心角的弧度, 如图 2-1 所示.

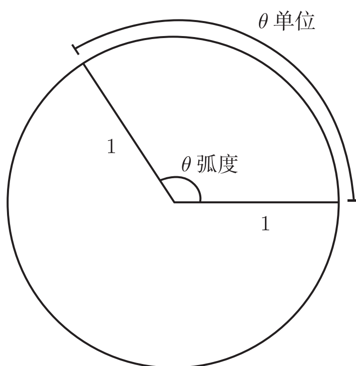


图　2-1

上图表示了一般情况, 但要紧的还是一些常用角的度和弧度表达. 首先, 你应该确实掌握, $90°$ 和 $\pi/2$ 弧度是一样的. 类似地, $180°$ 和 π 弧度是一样的, $270°$ 和 $3\pi/2$ 弧度是一样的. 一旦掌握了这几个角, 就试着将图 2-2 中所有的角在度与弧度之间来回转换吧.

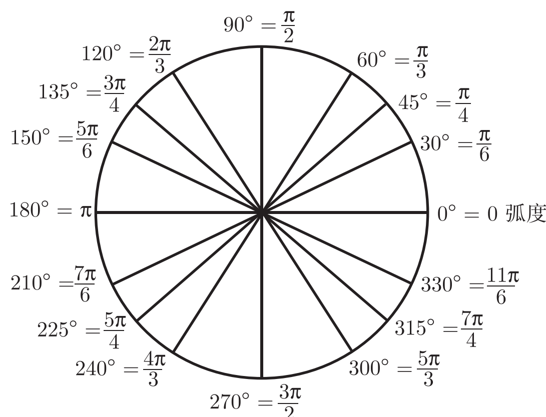


图　2-2

更一般地, 如果需要的话, 也可以使用公式

$$\text{用弧度度量的角} = \frac{\pi}{180} \times \text{用度度量的角}.$$



例如, 要想知道 $5\pi/12$ 弧度是多少度, 可求解

$$\frac{5\pi}{12} = \frac{\pi}{180} \times \text{用度计量的角},$$

你会发现 $5\pi/12$ 弧度就是 $(180/\pi) \times (5\pi/12) = 75°$. 事实上, 可以将弧度和度的转换看成是一种单位的转换, 如英里和公里的转换一样. 转换因数就是 π 弧度等于 180°.

到目前为止, 我们仅仅研究了角, 现在来看看三角函数吧. 显然, 你必须知道如何由三角形来定义三角函数. 假设我们有一个直角三角形, 除直角外的一角被记为 θ, 如图 2-3 所示. 那么, 基本公式为

$$\sin(\theta) = \frac{\text{对边}}{\text{斜边}}, \quad \cos(\theta) = \frac{\text{邻边}}{\text{斜边}}, \quad \tan(\theta) = \frac{\text{对边}}{\text{邻边}}.$$

当然, 如果变换了角 θ, 那么也必须变换其对边和邻边, 如图 2-4 所示. 毫不奇怪, 对边就是对着角 θ 的边, 而邻边则是挨着角 θ 的边. 不过, 斜边始终保持不变: 它是最长的那条边, 并始终对着直角.

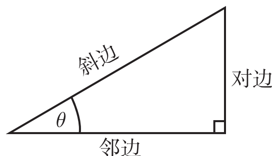


图　2-3

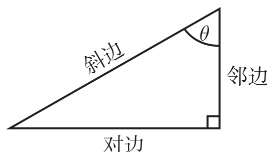


图　2-4

我们也会用到余割、正割和余切这些倒数函数, 它们的定义分别为

$$\csc(x) = \frac{1}{\sin(x)}, \quad \sec(x) = \frac{1}{\cos(x)}, \quad \cot(x) = \frac{1}{\tan(x)}.$$

如果你有计划要参加一次微积分的考试 (或者即便你没有), 我的一点建议是：请熟记常用角 $0, \pi/6, \pi/4, \pi/3, \pi/2$ 的三角函数值. 例如, 你能不假思索化简 $\sin(\pi/3)$ 吗？$\tan(\pi/4)$ 呢？如果你不能, 那么最好的情况下, 你通过画三角形来寻找答案, 从而白白浪费时间; 而最坏的情况下, 由于总是没有化简你的回答, 你白白丢掉分数. 解决的方法就是要熟记下表.



	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
sin	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
cos	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
tan	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	$\star$

表中的星号表示 $\tan(\pi/2)$ 无定义. 事实上, 正切函数在 $\pi/2$ 处有一条垂直渐近线 (从图像上看会很清楚, 我们将在 2.3 节对此进行研究). 无论如何, 你必须能够熟练地说出该表中的任意一项, 而且来回都要掌握! 这意味着你必须能够回答两类问题. 这两类问题的例子是：

(1) $\sin(\pi/3)$ 是什么？(使用该表, 答案是 $\sqrt{3}/2$.)

(2) 介于 0 到 $\pi/2$, 其正弦值为 $\sqrt{3}/2$ 的角是什么？(显然, 答案是 $\pi/3$.)

当然, 你必须能够回答该表中的每一项所对应的这两类问题. 就算我求大家了, 请背熟这张表! 数学不是死记硬背, 但有些内容是值得记忆的, 而这张表一定位列其中. 因此, 无论是制作记忆卡片, 让你的朋友来测验你, 还是每天抽一分钟记忆, 不管用什么办法, 请背熟这张表.

2.2 扩展三角函数定义域

上表 (你背熟了吗？) 仅仅包括介于 0 到 $\pi/2$ 的一些角. 但事实上, 我们可以取任意角的正弦或者余弦, 哪怕这个角是负的. 对于正切函数, 我们则不得不小心些. 例如, 上面我们看到的 $\tan(\pi/2)$ 是无定义的. 尽管如此, 我们还是能够对几乎每一个角取正切.

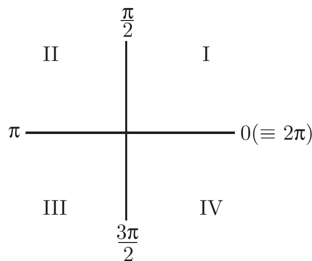


图 2-5

让我们首先来看看介于 0 到 2π (记住, 2π 就是 $360°$) 的角吧. 假设你想要计算 $\sin(\theta)$ (或 $\cos(\theta)$ 或 $\tan(\theta)$), 其中 θ 是介于 0 到 $\pi/2$ 的角. 为了看得更清楚, 我们先来画一个带有一点古怪标记的坐标平面, 如图 2-5 所示.

注意到坐标轴将平面分成了四个象限, 标记为 I 到 IV, 且标记的走向为逆时针

方向. 这些象限分别被称为第一象限、第二象限、第三象限和第四象限. 下一步是要画一条始于原点的射线 (就是半直线). 那么究竟是哪一条射线呢? 这取决于角 θ. 来想象一下, 你自己站在原点上, 面向 x 轴的正半轴. 现在沿着逆时针方向转动角 θ, 然后你沿着一条直线向前走. 你的足迹就是你要找的那条射线了.

现在, 图 2-5 (以及图 2-2) 中的其他标记就说得通了. 事实上, 如果你转动了角 $\pi/2$, 你将正面向上并且你的足迹将是 y 轴的正半轴. 如果你转动了角 π, 你将得到 x 轴的负半轴. 如果你转动了角 $3\pi/2$, 你将得到 y 轴的负半轴. 最后, 如果你转动了角 2π, 那么就又会回到了你起始的那个位置, 即面向 x 轴的正半轴. 这就好像你根本没转动过! 这就是为什么图中会有 $0 \equiv 2\pi$. 对于角度而言, 0 和 2π 是等价的.

好了, 让我们取某个角 θ 并以恰当的方式画出它. 或许它就在第三象限的某个地方, 如图 2-6 所示.

注意到我们将这条射线标记为 θ, 而不是这个角本身. 不管怎样, 现在在这条射线上选取某个点并从该点画一条垂线至 x 轴. 我们对三个量感兴趣：该点的 x 坐标和 y 坐标 (当然它们被称为 x 和 y), 以及该点到原点的距离, 我们称为 r. 注意, x 和 y 可能会同时为负 (事实上, 在图 2-7 中它们均为负). 然而, r 总是正的, 因为它是距离. 事实上, 根据毕达哥拉斯定理 (即勾股定理), 不管 x 和 y 是正还是负, 我们总会有 $r=\sqrt{x^2+y^2}$. (平方会消除任何负号.)

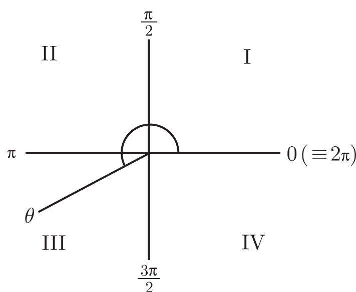


图　2-6

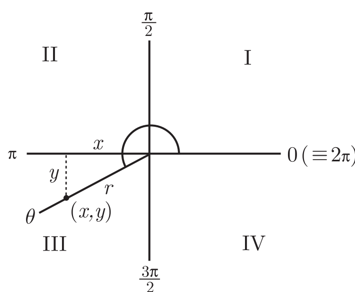


图　2-7

有了这三个量, 我们就可以定义如下的三个三角函数了：

$$\sin(\theta)=\frac{y}{r},\quad \cos(\theta)=\frac{x}{r},\quad \tan(\theta)=\frac{y}{x}.$$

将量 x、y 和 r 分别解释为邻边、对边和斜边, 这些函数恰好就是 2.1 节中的固定公式了. 不过等一下, 如果你在那条射线上选取了另外一个点, 那会是什么样子呢? 这不要紧, 因为你得到的新的三角形和原来的那个三角形是相似的, 而上述比值不会受到任何影响. 事实上, 为方便起见, 我们常常假设 $r=1$, 这样得到的点 $(x,\ y)$ 会落在所谓的**单位圆** (就是以原点为中心, 半径为 1 的圆) 上.



现在来看一个例子. 假设, 我们想求 $\sin(7\pi/6)$. 首先, $7\pi/6$ 会在第几象限呢?

我们需要决定 $7\pi/6$ 会出现在列表 $0, \pi/2, \pi, 3\pi/2, 2\pi$ 的哪个地方. 事实上, $7/6$ 大于 1 但小于 $3/2$, 故 $7\pi/6$ 在 π 和 $3\pi/2$ 之间. 事实上, 图 2-8 看起来很像前面的例子.

因此, 角 $7\pi/6$ 在第三象限. 然后, 我们选取了该射线上的一点, 该点至原点的距离 $r = 1$, 并从该点至 x 轴做了一条垂线. 由前述公式可知, $\sin(\theta) = y/r = y$ (因为 $r = 1$), 因此, 我们还是要求出 y. 好吧, 那个小角, 就是在 $7\pi/6$ 处的射线和 x 轴的负半轴 (其为 π) 之间的角一定是这两个角的差, 即 $\pi/6$. 这个小角被称为**参考角**. 一般来说, θ 的参考角是在表示角 θ 的射线和 x 轴之间的最小的角, 它必定介于 0 到 $\pi/2$. 在我们的例子中, 到 x 轴的最短路径是向上, 所以参考角如图 2-9 所示. 因此, 在那个小三角形中, 我们知道 $r = 1$, 以及角为 $\pi/6$. 似乎答案就是 $y = \sin(\pi/6) = 1/2$, 但这是错的! 由于在 x 轴的下方, y 一定为负值. 也就是说, $y = -1/2$. 因为 $\sin(\theta) = y$, 我们也就证明了 $\sin(7\pi/6) = -1/2$. 对于余弦来说, 也可以重复这个过程, 求出 $x = -\cos(\pi/6) = -\sqrt{3}/2$. 毕竟, 由于点 (x, y) 在 y 轴的左侧, 因此 x 必须为负. 这样就证明了 $\cos(7\pi/6) = -\sqrt{3}/2$, 并且识别出点 (x, y) 即为点 $\left(-\sqrt{3}/2, -1/2\right)$.

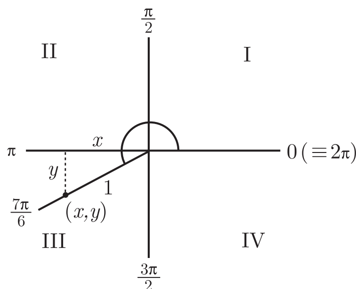


图 2-8

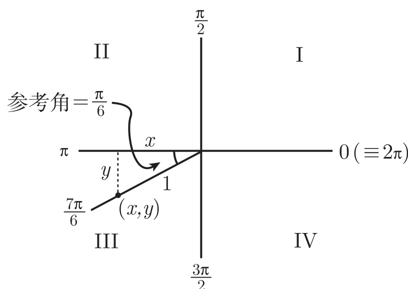


图 2-9

2.2.1 ASTC 方法

上例中的关键是将 $\sin(7\pi/6)$ 和 $\sin(\pi/6)$ 联系起来, 其中 $\pi/6$ 是 $7\pi/6$ 的参考角. 事实上, 并不难看出任意角的正弦就是其参考角正弦的正值或负值! 这就使问题缩小到两种可能性上, 而且没有必要再纠缠于 x, y 或 r 如此这般麻烦. 因此, 在我们的例子中, 只需要求出 $7\pi/6$ 的参考角, 即 $\pi/6$; 这就会立即可知 $\sin(7\pi/6)$ 等于 $\sin(\pi/6)$ 或 $-\sin(\pi/6)$, 而我们只需从中选出正确的结果. 我们发现, 结果是负的那个, 因为 y 是负的.

事实上, 在第三或第四象限中的任意角的正弦必定为负, 因为那里的 y 为负. 类似地, 在第二或第三象限中的任意角的余弦必定为负, 因为那里的 x 为负. 正切是比值 y/x, 它在第二和第四象限为负 (由于 x 和 y 中的一个为负, 但不全为负), 而在第一和第三象限为正.

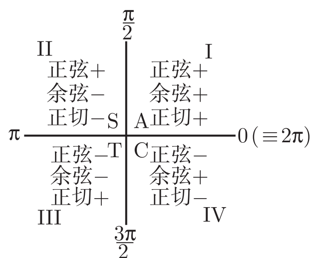


图　2-10

我们来总结一下这些发现. 首先, 所有三个函数在第一象限 (I) 中均为正. 在第二象限 (II) 中, 只有正弦为正, 其他两个函数均为负. 在第三象限 (III) 中, 只有正切为正, 其他两个函数均为负. 最后, 在第四象限 (IV) 中, 只有余弦为正, 其他两个函数均为负. 具体如图 2-10 所示.

事实上, 你只需要记住图表中的字母 ASTC 就行了. 它们会告诉你在那个象限中哪个函数为正. “A” 代表 “全部”, 意味着所有的函数在第一象限均为正. 显然, 其余的字母分别代表正弦、正切和余弦. 在我们的例子中, $7\pi/6$ 在第三象限, 所以只有正切函数在那里为正. 特别地, 正弦函数为负, 又由于我们已经把 $\sin(7\pi/6)$ 的可能取值缩小到 $1/2$ 或 $-1/2$ 了, 因此结果一定是负的那个, 即 $\sin(7\pi/6) = -1/2$.

ASTC 图唯一的问题在于, 它没有告诉我们该如何处理角 0, $\pi/2$, π 或 $3\pi/2$, 因为它们都位于坐标轴上. 这种情况下, 最好是先忘记所有 ASTC 的内容, 然后以恰当的方式画一个 $y = \sin(x)$ (或 $\cos(x)$, 或 $\tan(x)$) 的图像, 并且从图像中读取数值. 我们将在 2.3 节对此进行研究.

以下是用 ASTC 方法来求介于 0 到 2π 的角的三角函数值的总结.

(1) 画出象限图, 确定在该图中你感兴趣的角在哪里, 然后在图中标出该角.

(2) 如果你想要的角在 x 轴或 y 轴上 (即没有在任何象限中), 那么就画出三角函数的图像, 从图像中读取数值 (2.3 节有一些例子).

(3) 否则, 找出在代表我们想要的那个角的射线和 x 轴之间最小的角, 这个角被称为参考角.

(4) 如果可以, 使用那张重要的表来求出参考角的三角函数值. 那就是你需要的答案, 除了你可能还需要在得到的值前面添一个负号.

(5) 使用 ASTC 图来决定你是否需要添一个负号.

来看一些例子. 如何求 $\cos(7\pi/4)$ 和 $\tan(9\pi/13)$ 呢? 我们一个一个地看. 对于 $\cos(7\pi/4)$, 我们注意到 $7/4$ 介于 $3/2$ 和 2 之间, 故该角必在第四象限, 如图 2-11 所示.

为了求出参考角, 注意到我们必须向上走到 2π (注意! 不是到 0), 因此, 参考角就是 2π 和 $7\pi/4$ 的差, 即 $(2\pi - 7\pi/4)$, 或简化为 $\pi/4$. 所以 $\cos(7\pi/4)$ 是正的或负的 $\cos(\pi/4)$. 根据表, $\cos(\pi/4)$ 是 $1/\sqrt{2}$. 但到底是正的还是负的呢? 由 ASTC 图可知, 在第四象限中余弦为正, 故结果为正的那个: $\cos(7\pi/4) = 1/\sqrt{2}$.

现在来看一下 $\tan(9\pi/13)$. 我们发现 $9/13$ 介于 $1/2$ 和 1 之间, 故角 $9\pi/13$ 在第二象限, 如图 2-12 所示.

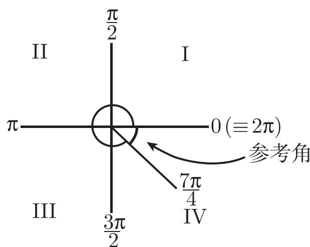


图 2-11

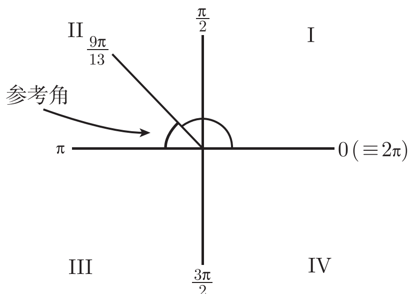


图 2-12

这一次, 我们需要走到 π 以到达 x 轴, 故参考角就是 π 和 9π/13 的差, 即 (π−9π/13), 或简化为 4π/13. 这样, 我们知道 tan (9π/13) 是正的或负的 tan (4π/13). 哎呀, 可是数 4π/13 没有在我们的表里面, 因此不能化简 tan (4π/13). 可我们还是需要确定它是正的还是负的. 那好, ASTC 图显示, 在第二象限中只有正弦为正, 故正切一定为负, 于是 tan (9π/13) = −tan (4π/13). 这就是不使用近似可以得到的最简形式. 在求解微积分问题的时候, 我不建议取近似结果, 除非题目中有明确要求. 一个常见的误解是, 当你计算如同 −tan (4π/13) 这样的问题时, 由计算器计算出来的数就是正确答案. 其实, 那只是一个近似! 所以你不应该写

$$-\tan(4\pi/13) = -1.448\,750\,113,$$

因为它不正确. 就应该写 −tan (4π/13), 除非有特别的要求, 让做近似. 在那种情况下, 使用约等号和更少的小数位数, 并恰当化整近似 (除非要求保留更多小数位数):

$$-\tan(4\pi/13) \approx -1.449.$$

顺便说一下, 你应该少用计算器. 事实上, 一些大学甚至不允许在考试中使用计算器! 因此, 你应该尽量避免使用计算器.

2.2.2 [0, 2π] 以外的三角函数

还有一个问题, 就是如何取大于 2π 或小于 0 的角的三角函数. 事实上, 这并不太难, 简单地加上或减去 2π 的倍数, 直到你得到的角在 0 和 2π 之间. 你看, 它并不是在 2π 就完了. 它是一直在旋转. 例如, 如果我让你站在一点面向正东, 然后逆时针方向旋转 450°, 一种自然的做法是, 你旋转一整周, 然后再旋转 90°. 现在你应该是面向正北. 当然, 另一种不那么头晕目眩的做法是, 你只逆时针方向旋转 90°, 而你面向的是同样的方向. 因此, 450° 和 90° 是等价的角. 当然, 这对于弧度来说也一样. 这种情况下, 5π/2 弧度和 π/2 弧度是等价的角. 但为什么要止步于旋转一周呢? 9π/2 弧度又如何? 这和旋转 2π 两次 (这样我们得到 4π), 然后再旋转 π/2 是

一样的. 因此, 在得到最终的 $\pi/2$ 之前, 我们做了两周徒劳的旋转. 旋转周数无关紧要, 我们再次得到 $9\pi/2$ 和 $\pi/2$ 等价. 这个过程可以被无限地扩展下去, 以得到等价于 $\pi/2$ 的角的一个家族：

$$\frac{\pi}{2}, \frac{5\pi}{2}, \frac{9\pi}{2}, \frac{13\pi}{2}, \frac{17\pi}{2}, \cdots.$$

当然, 这其中的每一个角都比前一个角多一个整周旋转, 即 2π. 但这仍然还没算完. 如果你做了所有这些逆时针旋转, 并感到头晕目眩, 或许你也会要求做一个或两个顺时针旋转来缓和一下. 这就相当于一个负角. 特别地, 如果你面向东, 我让你逆时针旋转 $-270°$, 对我这个怪异要求唯一合理的解释就是顺时针旋转 $270°$ (或 $3\pi/2$). 显然, 你最终仍然会面向正北, 因此, $-270°$ 和 $90°$ 一定是等价的. 确实, 我们将 $360°$ 加到 $-270°$ 上就会得到 $90°$. 使用弧度, 我们则看到, $-3\pi/2$ 和 $\pi/2$ 是等价的角. 另外, 我们可以要求更多负的 (顺时针方向) 整周旋转. 最后, 以下就是等价于 $\pi/2$ 的角的完全的集合：

$$\cdots, -\frac{15\pi}{2}, -\frac{11\pi}{2}, -\frac{7\pi}{2}, -\frac{3\pi}{2}, \frac{\pi}{2}, \frac{5\pi}{2}, \frac{9\pi}{2}, \frac{13\pi}{2}, \frac{17\pi}{2}, \cdots.$$

这个序列没有开端也没有结束. 当我说它是 "完全的" 时, 我用前后两头的省略号代表了无穷多个角. 为了避免这些省略号, 我们可以使用集合符号 $\{\pi/2 + 2\pi n\}$, 其中 n 可以取所有整数.

来看一下是否可以应用它吧. 如何求 $\sec(15\pi/4)$ 呢? 首先, 注意到如果我们能够求出 $\cos(15\pi/4)$, 所要做的就是取其倒数以得到 $\sec(15\pi/4)$. 因此, 让我们先求 $\cos(15\pi/4)$. 由于 $15/4$ 大于 2, 让我们先试着消去 2. 这样, $15/4 - 2 = 7/4$, 现在它介于 0 和 2 之间, 这看上去很有希望了. 代入 π, 我们看到 $\cos(15\pi/4)$ 和 $\cos(7\pi/4)$ 是一样的, 并且我们已经求出其结果为 $1/\sqrt{2}$. 因此, $\cos(15\pi/4) = 1/\sqrt{2}$. 取其倒数, 我们发现 $\sec(15\pi/4)$ 就是 $\sqrt{2}$.

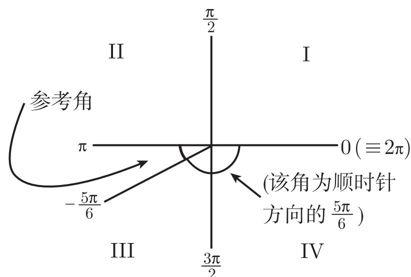


图　2-13

最后, $\sin(-5\pi/6)$ 又如何呢? 有很多方法来求解此问题, 但上面提到的方法是试着将 2π 的倍数加到 $-5\pi/6$ 上, 直到结果是介于 0 到 2π 的. 事实上, 2π 加上 $-5\pi/6$ 得 $7\pi/6$, 因此, $\sin(-5\pi/6) = \sin(7\pi/6)$, 后者我们已经知道等于 $-1/2$. 另外, 我们也可以直接画图 2-13.

现在, 你必须找出图中的参考角. 不难看出, 它是 $\pi/6$, 然后一如前述.

2.3 三角函数的图像

记住正弦、余弦和正切函数的图像会非常有用. 这些函数都是**周期的**, 这意味着, 它们从左到右反复地重复自己. 例如, 我们考虑 $y = \sin(x)$. 从 0 到 2π 的图像看上去如图 2-14 所示.

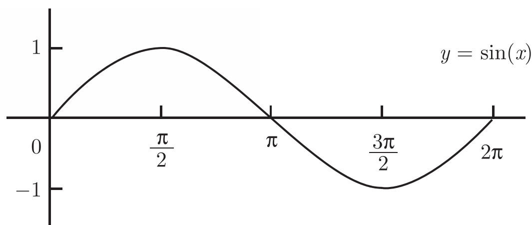


图 2-14

你应该做到能够不假思索就画出这个图像, 包括 $0, \pi/2, \pi, 3\pi/2$ 和 2π 的位置. 由于 $\sin(x)$ 以 2π 为单位重复 (我们说 $\sin(x)$ 是 x 的周期函数, 其周期为 2π), 通过重复该模式, 我们可以对图像进行扩展, 得到图 2-15.

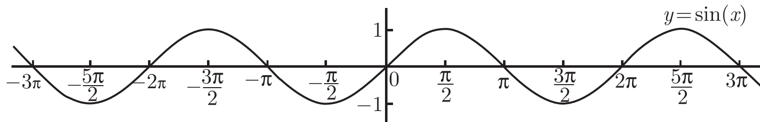


图 2-15

从图像中读值, 可以看到 $\sin(3\pi/2) = -1$, $\sin(-\pi) = 0$. 正如之前注意到的, 你应该这样去处理 $\pi/2$ 的倍数的问题, 而不用再找参考角那么麻烦了. 另一个值得注意的是, 该图像关于原点有 180° 点对称性, 这意味着, $\sin(x)$ 是 x 的奇函数. (我们在 1.4 节中分析过奇偶函数.)

$y = \cos(x)$ 的图像和 $y = \sin(x)$ 的图像类似. 当 x 在从 0 到 2π 上变化时, 它看起来就像图 2-16.

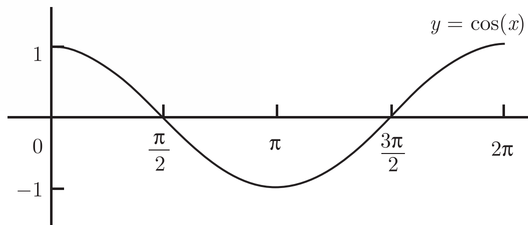


图 2-16

现在, 利用 $\cos(x)$ 是周期函数及其周期为 2π 这一事实, 可对该图像进行扩展, 得到图 2-17.

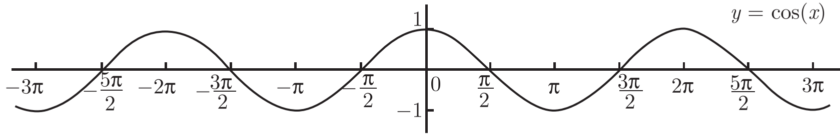


图　2-17

例如, 如果你想要求 $\cos(\pi)$, 只需从图像上读取, 你会看到结果是 -1. 此外, 注意到该图像关于 y 轴有镜面对称性. 这说明, $\cos(x)$ 是 x 的偶函数.

现在, $y = \tan(x)$ 略有不同. 最好是先画出 x 介于 $-\pi/2$ 到 $\pi/2$ 的图像, 如图 2-18 所示.

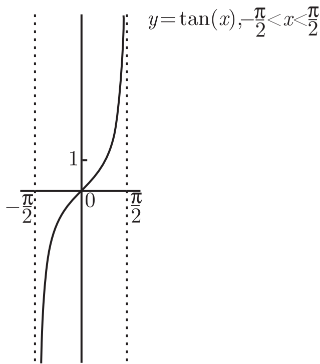


图　2-18

与正弦函数和余弦函数不同的是, 正切函数有垂直渐近线. 此外, 它的周期是 π, 而不是 2π. 因此, 上述图样可以被重复以便得到 $y = \tan(x)$ 的全部图像, 如图 2-19 所示.

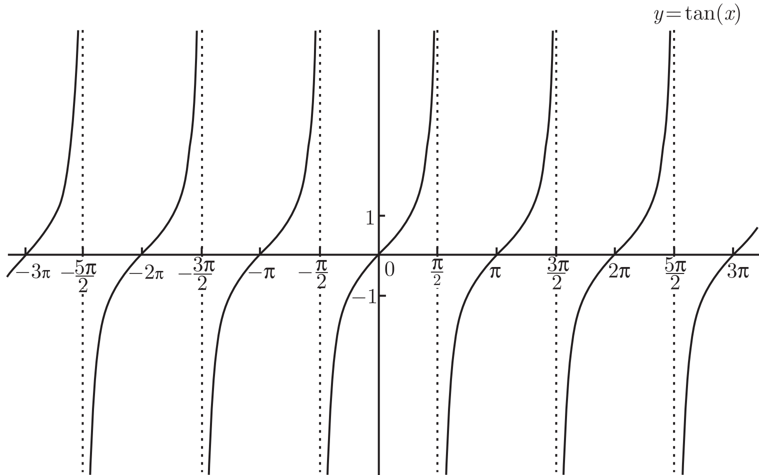


图　2-19

很明显, 当 x 是 $\pi/2$ 的奇数倍数时, $y=\tan(x)$ 有垂直渐近线 (因而此处是无定义的). 此外, 图像的对称性表明, $\tan(x)$ 是 x 的奇函数.

$y=\sec(x)$、$y=\csc(x)$ 及 $y=\cot(x)$ 的函数图像也值得我们去学习, 它们分别如图 2-20、图 2-21 及图 2-22 所示.

从它们的图像中, 可以得到所有六个基本三角函数的对称性的性质, 这些也都值得学习.

> $\sin(x)$、$\tan(x)$、$\cot(x)$, 及 $\csc(x)$ 都是 x 的奇函数. $\cos(x)$ 和 $\sec(x)$ 都是 x 的偶函数.

因此, 对于所有的实数 x, 我们有 $\sin(-x)=-\sin(x)$, $\tan(-x)=-\tan(x)$, $\cos(-x)=\cos(x)$.

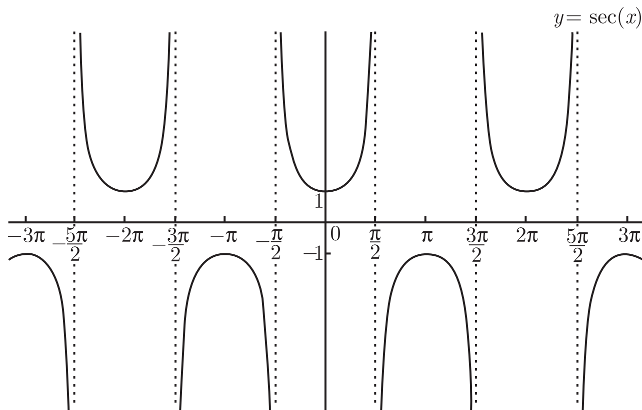


图 2-20

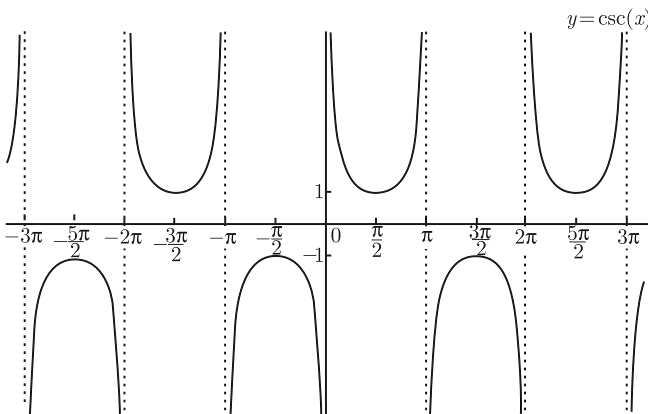


图 2-21

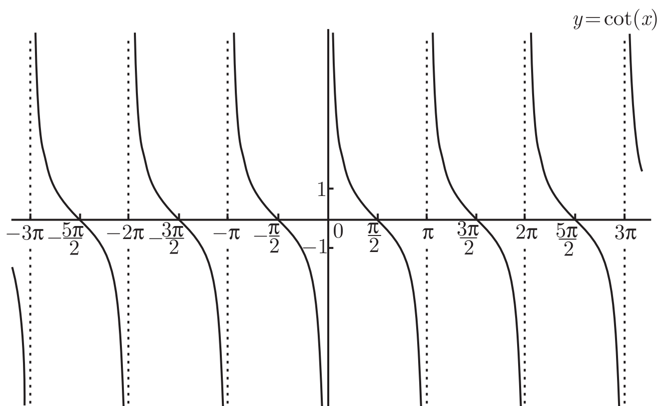


图　2-22

2.4　三角恒等式

三角函数间的关系用来十分方便. 首先, 注意到正切和余切可以由正弦和余弦来表示:

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)}.$$

(有时, 根据这些恒等式, 用正弦和余弦来代替每一个正切和余切会有帮助, 但这只是你被卡住时不得已而为之的下下策.)

所有三角恒等式中最重要的就是毕达哥拉斯定理了 (用三角函数表示),

$$\boxed{\cos^2(x) + \sin^2(x) = 1.}$$

这对于任意的 x 都成立. (为什么这是毕达哥拉斯定理呢? 如果直角三角形的斜边是 1, 其中一个角为 x, 自己验证三角形的其他两条边长就是 $\cos(x)$ 和 $\sin(x)$.)



现在, 让这个等式两边同除以 $\cos^2(x)$. 你应该能够得到以下结果:

$$\boxed{1 + \tan^2(x) = \sec^2(x).}$$

该公式在微积分里也会经常出现. 另外, 你也可以将毕达哥拉斯定理等式两边同除以 $\sin^2(x)$, 得到以下等式:

$$\boxed{\cot^2(x) + 1 = \csc^2(x).}$$

这个公式好像没有其他公式出现得那么频繁.

三角函数之间还有其他一些关系. 你注意到了吗? 一些函数的名字是以音节 “co” 开头的. 这是 “互余” (complementary) 的简称. 说两个角互余, 意味着它们的和是 $\pi/2$ (或 $90°$). 可不是说它们相互恭维. 好吧, 不玩双关了, 事实是有以下一般关系:

$$\text{三角函数}\,(x) = \text{co-三角函数}\left(\frac{\pi}{2} - x\right).$$

特别地, 有：

$$\sin(x) = \cos\left(\frac{\pi}{2} - x\right), \quad \tan(x) = \cot\left(\frac{\pi}{2} - x\right), \quad \sec(x) = \csc\left(\frac{\pi}{2} - x\right).$$

甚至当三角函数名中已经带有一个 “co” 时, 以上公式仍然适用; 你只需要认识到, 余角的余角就是原始的角! 例如, co-co-sin 事实上就是 sin, co-co-tan 事实上就是 tan. 基本上, 这意味着我们还可以说：

$$\cos(x) = \sin\left(\frac{\pi}{2} - x\right), \quad \cot(x) = \tan\left(\frac{\pi}{2} - x\right), \quad \csc(x) = \sec\left(\frac{\pi}{2} - x\right).$$

最后, 还有一组恒等式值得我们学习. 这些恒等式涉及角的和与倍角公式. 特别地, 我们应该记住下列公式：

$$\boxed{\begin{aligned}\sin(A+B) &= \sin(A)\cos(B) + \cos(A)\sin(B)\\ \cos(A+B) &= \cos(A)\cos(B) - \sin(A)\sin(B).\end{aligned}}$$

还应该记住, 你可以切换所有的正号和负号, 得到一些相关的公式：

$$\sin(A-B) = \sin(A)\cos(B) - \cos(A)\sin(B)$$
$$\cos(A-B) = \cos(A)\cos(B) + \sin(A)\sin(B).$$

对于上述方框公式中的 $\sin(A+B)$ 和 $\cos(A+B)$, 令 $A = B = x$, 我们就会得到另一个有用的结果. 很明显, 正弦公式是 $\sin(2x) = 2\sin(x)\cos(x)$. 但让我们更仔细看一下余弦公式. 它会变成 $\cos(2x) = \cos^2(x) - \sin^2(x)$; 这本身没错, 但更有用的是使用毕达哥拉斯定理 $\sin^2(x) + \cos^2(x) = 1$ 将 $\cos(2x)$ 表示成为 $2\cos^2(x) - 1$ 或 $1 - 2\sin^2(x)$ (自己验证一下它们是成立的!). 综上, 倍角公式为：

$$\boxed{\begin{aligned}&\sin(2x) = 2\sin(x)\cos(x)\\ &\cos(2x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x).\end{aligned}}$$



那么如何用 $\sin(x)$ 和 $\cos(x)$ 来表示 $\sin(4x)$ 呢? 我们可以将 $4x$ 看作两倍的 $2x$, 并使用正弦恒等式, 写作 $\sin(4x) = 2\sin(2x)\cos(2x)$. 然后应用两个恒等式, 得到

$$\sin(4x) = 2(2\sin(x)\cos(x))(2\cos^2(x) - 1) = 8\sin(x)\cos^3(x) - 4\sin(x)\cos(x).$$

类似地,

$$\cos(4x) = 2\cos^2(2x) - 1 = 2(2\cos^2(x) - 1)^2 - 1 = 8\cos^4(x) - 8\cos^2(x) + 1.$$

你不用记这后两个公式; 相反, 你要确保理解了如何使用倍角公式来推导它们.

如果你能够掌握本章涉及的所有三角学内容, 就能够很好地学习本书的剩余部分了. 因此, 抓紧时间消化这些知识吧. 做一些例题, 并确保你记住了那张很重要的表格和所有方框公式.

第 3 章　极 限 导 论

如果没有极限的概念, 那么微积分将不复存在. 这意味着, 我们将用大量的时间来研究它们. 事实证明, 虽然恰当地定义一个极限是件相当棘手的事情, 但你仍然有可能对极限有个直观理解, 而无须深入其中的具体细节. 这对于解决微分和积分问题已经足够了. 因此, 本章仅仅包含对极限的直观描述; 正式描述请参见附录 A. 总的来说, 以下就是我们会在本章讲解的内容:

- 对于极限是什么的直观概念;
- 左、右与双侧极限, 及在 ∞ 和 $-\infty$ 处的极限;
- 何时极限不存在;
- 三明治定理 (也称作 "夹逼定理").

3.1　极限: 基本思想

让我们开始吧. 我们从某个函数 f 和 x 轴上的一点出发, 该点称为 a. 需要理解的是: 当 x 非常非常接近于 a, 但不等于 a 时, $f(x)$ 是什么样子的? 这是一个非常奇怪的问题, 人类相对晚近才发展出微积分很可能就是因为这个原因吧.

这里有一个例子, 说明了为什么要提出这样的问题. 令 f 的定义域为 $\mathbb{R}\setminus\{2\}$ (除 2 以外的所有实数), 并设 $f(x) = x - 1$. 这可以写作:

$$f(x) = x - 1 \quad 当\ x \neq 2.$$

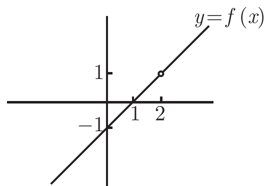


图　3-1

这看起来好像是一个古怪的函数. 毕竟, 到底为什么要将 2 从定义域中去除掉呢? 其实, 在下一章就会看到, f 很自然地就是个有理函数 (参见 4.1 节的第二个例子) 不过现在, 让我们姑且接受 f 的定义, 并画出其图像, 如图 3-1 所示.

那么 $f(2)$ 是什么呢? 或许你会说 $f(2) = 1$, 但这是大错特错了, 因为 2 根本不在 f 的定义域中. 你所能给出的最好回答就是 $f(2)$ 是无定义的. 另一方面, 当 x 非常非常接近于 2 的时候, 我们可以找到一些 $f(x)$ 的值, 并看看将会有什么发生. 例如, $f(2.01) = 1.01$, $f(1.999) = 0.999$. 稍作思考, 你会发现当 x 非常非常接近于 2 的时候, $f(x)$ 的值会非常非常接近于 1.

还有, 只要令 x 充分地接近于 2, 那么你想多接近于 1 就能多接近于 1, 却又不是真的达到 1. 例如, 如果你想要 $f(x)$ 在 1 ± 0.0001 内, 可以取在 1.9999 和

2.0001 之间的任意的 x 值 (当然, 除了 $x=2$, 这是禁止的). 如果你想要 $f(x)$ 在 $1\pm0.000\,007$ 内, 那么选取 x 的时候, 你不得不更细心一点. 这一次, 你需要取在 1.999 993 和 2.000 007 之间的任意值了 (当然, 还是除了 2).

这些思想会在附录 A 的 A.1 节里有更详细的描述. 不过现在, 让我们回到正题, 直接写出

$$\lim_{x\to2}f(x)=1.$$

如果你大声将它读出来, 它听起来应该像是 "当 x 趋于 2, $f(x)$ 的极限等于 1". 再次说明, 这意味着, 当 x 接近于 2(但不等于 2) 时, $f(x)$ 的值接近于 1. 那到底有多近呢? 你想要多近就能多近. 以上陈述的另外一个写法是

$$f(x)\to1 \text{ 当 } x\to2.$$

这个写法更难用来计算, 但其意义很清晰: 当 x 沿着数轴从左侧或者从右侧趋近于 2 时, $f(x)$ 的值会非常非常接近于 1(并保持接近的状态!).

现在, 取上述函数 f 并对它做一点改动. 假设有一个新的函数 g, 其图像如图 3-2 所示.

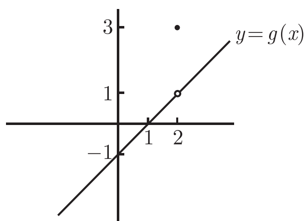


图 3-2

函数 g 的定义域是所有实数, 并且 $g(x)$ 可以被定义为如下的分段函数:

$$g(x)=\begin{cases}x-1 & \text{如果 } x\neq2,\\ 3 & \text{如果 } x=2\end{cases}$$

$\lim\limits_{x\to2}g(x)$ 是什么呢? 这里的关键是, $g(2)$ 的值和该极限是不相关的! 只有那些在 x 接近于 2 时的 $g(x)$ 的值, 而不是在 2 处的值, 才是问题的关键. 如果忽略 $x=2$, 函数 g 和之前的函数 f 就是完全相同的. 因此, 尽管 $g(2)=3$, 我们还是有 $\lim\limits_{x\to2}g(x)=1$.

这里的要点是, 当你写出

$$\lim_{x\to2}f(x)=1,$$

的时候, 等式左边实际上不是 x 的函数! 要记住, 以上等式是说, 当 x 接近于 2 时, $f(x)$ 接近于 1. 事实上, 我们可以将 x 替换成其他任意字母, 上式仍然成立. 例如, 当 q 接近于 2 时, $f(q)$ 接近于 1, 因此我们有

$$\lim_{q\to2}f(q)=1.$$

也可以写成

$$\lim_{b\to2}f(b)=1,\quad \lim_{z\to2}f(z)=1,\quad \lim_{\alpha\to2}f(\alpha)=1,$$

如此等等, 直到用光了所有的字母和符号! 这里的要点是, 在极限

$$\lim_{x\to2}f(x)=1,$$

中, 变量 x 只是一个**虚拟变量**. 它是一个暂时的标记, 用来表示某个 (在上述情况下) 非常接近于 2 的量. 它可以被替换成其他任意字母, 只要替换是彻底的; 同样,

当你求出极限的值时, 结果不可能包含这个虚拟变量. 所以对虚拟变量你要灵活处理.

3.2　左极限与右极限

我们已经看到, 极限描述了函数在一个定点附近的行为. 现在想想看, 你会如何描述图 3-3 中 $h(x)$ 在 $x=3$ 附近的行为.

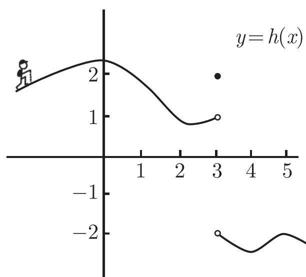


图　3-3

当然, 就趋于极限的行为而言, $h(3)=2$ 实际上是无关紧要的. 现在, 当你从左侧接近于 $x=3$ 时会发生什么呢? 想象一下, 你是图中的远足者, 顺着山势上下. $h(x)$ 的值会告诉你, 当你的水平位置是 x 时, 你所在高度是多少. 因此, 如果你从图的左边向右走, 那么当你的水平位置接近于 3 时, 你所在高度就会接近于 1. 当然, 当到达 $x=3$ 时你会陡然坠落 (更不用说那个古怪的小突起), 但暂时我们不关心. 这时任何在 $x=3$ 右侧的值, 包含 $x=3$ 本身对应的值, 都是无关紧要的. 因此, 就可以看到 $h(x)$ 在 $x=3$ 的**左极限**等于 1.

另一方面, 如果你从图的右边向左走, 那么当你的水平位置接近于 $x=3$ 时, 你所在高度就会接近于 -2. 这就是说, $h(x)$ 在 $x=3$ 的**右极限**等于 -2. 这时任何在 $x=3$ 左侧 (包含 $x=3$ 本身) 的值都是无关紧要的.

可将上述发现总结如下:

$$\lim_{x\to 3^-} h(x)=1 \quad 及 \quad \lim_{x\to 3^+} h(x)=-2.$$

在上面第一个极限中 3 后的小减号表示该极限是一个左极限, 第二个极限中 3 后的小加号表示该极限是一个右极限. 要在 3 的后面写上减号或加号, 而不是在前面, 这是非常重要的! 例如, 如果你写成

$$\lim_{x\to -3} h(x),$$

那么指的是 $h(x)$ 在 $x=-3$ 时的通常的双侧极限, 而不是 $h(x)$ 在 $x=3$ 时的左极限. 这确实是两个完全不同的概念. 顺便说一下, 在左极限的极限符号底下写 $x\to 3^-$ 的理由是, 此极限只涉及小于 3 的 x 的值. 也就是说, 你需要在 3 上减一点点来看会有什么情况发生. 类似地, 对于右极限, 当你写 $x\to 3^+$ 的时候, 这意味着你只需要考虑如果在 3 上加一点点会有什么情况发生.

正如我们将在下一节看到的, 极限不是总存在的. 但这里的要点是: 通常的双侧极限在 $x=a$ 处存在, 仅当左极限和右极限在 $x=a$ 处都存在且相等! 在这种情况下, 这三个极限 (双侧极限、左极限和右极限) 都是一样的. 用数学的语言描述, 我们说,

$$\lim_{x\to a^-} f(x) = L \text{ 且 } \lim_{x\to a^+} f(x) = L$$

等价于

$$\lim_{x\to a} f(x) = L.$$

如果左极限和右极限不相等, 例如上述例子中的函数 h, 那么双侧极限不存在. 我们写作

$$\lim_{x\to 3} h(x) \text{ 不存在}$$

或使用缩写 "DNE" 表示 "不存在".

3.3 何时不存在极限

我们刚刚看到, 当相应的左极限和右极限不相等时双侧极限不存在. 这里有一个更戏剧性的例子. 考虑 $f(x) = 1/x$ 的图像, 如图 3-4 所示. $\lim\limits_{x\to 0} f(x)$ 是什么呢? 双侧极限在那里不大可能存在. 因此, 我们先来试着求一下右极限, $\lim\limits_{x\to 0^+} f(x)$. 看一下图像, 当 x 是正的且接近于 0 时, $f(x)$ 看起来好像非常大. 特别是, 当 x 从右侧滑向 0 时, 它看起来并不接近于任何数; 它就是变得越来越大了. 但会有多大呢? 它会比你能想象到的任何数都大! 我们说该极限是无穷大, 并写作

$$\lim_{x\to 0^+} \frac{1}{x} = \infty.$$

类似地, 这里的左极限是 $-\infty$, 因为当 x 向 0 上升时, $f(x)$ 会变得越来越负. 这就是说,

$$\lim_{x\to 0^-} \frac{1}{x} = -\infty.$$

由于左极限和右极限不相等, 故双侧极限显然不存在. 另一方面, 考虑函数 g, 其定义为 $g(x) = 1/x^2$, 其图像如图 3-5 所示.

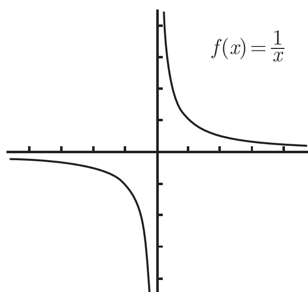


图 3-4

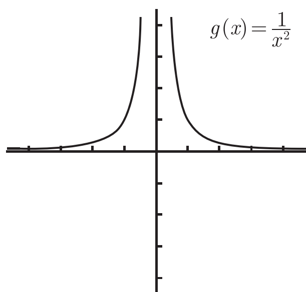


图 3-5

此函数在 $x = 0$ 处的左极限和右极限都是 ∞, 因此你也可以说 $\lim\limits_{x\to 0} 1/x^2 = \infty$. 顺便说一下, 现在我们有了一个关于 "垂直渐近线" 的正式定义:

> "f 在 $x=a$ 处有一条垂直渐近线" 说的是, $\lim\limits_{x\to a^+} f(x)$ 和 $\lim\limits_{x\to a^-} f(x)$, 其中至少有一个极限是 ∞ 或 $-\infty$.

现在, 可能会出现左极限或右极限不存在的情况吗？答案是肯定的! 例如, 让我们来看一个怪异的函数 g, 其定义为 $g(x)=\sin(1/x)$. 此函数的图像看起来会是什么样的呢？首先, 让我们来看一下 x 的正值. 由于 $\sin(x)$ 在 $x=\pi, 2\pi, 3\pi, \cdots$ 上的值全为 0, 因而 $\sin(1/x)$ 在 $1/x=\pi, 2\pi, 3\pi, \cdots$ 上的值全为 0. 我们取其倒数, 会发现 $\sin(1/x)$ 在 $x=\dfrac{1}{\pi}, \dfrac{1}{2\pi}, \dfrac{1}{3\pi}, \cdots$ 上的值全为 0. 这些数就是 $\sin(1/x)$ 的 x 轴截距. 在数轴上, 它们看起来如图 3-6 所示.

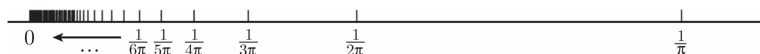


图　3-6

正如你看到的, 当接近于 0 的时候, 它们都挤在了一起. 由于在每一个 x 轴截距之间, $\sin(x)$ 向上走到 1 或向下走到 -1, 因此, $\sin(1/x)$ 也一样. 把目前已知的画出来, 可得到图 3-7.

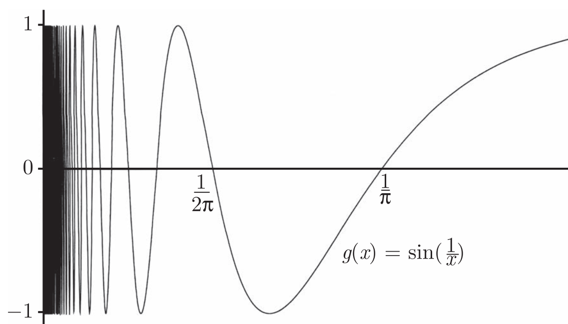


图　3-7

那么 $\lim\limits_{x\to 0^+} \sin(1/x)$ 是什么呢？以上图像在 $x=0$ 附近很杂乱. 它无限地在 1 和 -1 之间振荡, 当你从右侧向 $x=0$ 处移动时, 振荡会越来越快. 这里没有垂直渐近线, 也没有极限①. 当 x 从右侧趋于 $x=0$ 时, 该函数不趋于任何数. 因此可以说, $\lim\limits_{x\to 0^+} \sin(1/x)$ 不存在 (DNE). 我们会在下一节将 $y=\sin(1/x)$ 的图像补充完整.

3.4　在 ∞ 和 $-\infty$ 处的极限

还有一类需要研究的极限. 我们已经研究了在接近一点 $x=a$ 时的函数行为. 然而在有些情况下, 重要的是要理解当 x 变得非常大时, 一个函数的行为如何. 换

① 正式的证明请参见附录 A 的 A.3.4 节.

句话说, 我们感兴趣的是, 研究当变量 x 趋于 ∞ 时函数的行为. 我们想写出

$$\lim_{x\to\infty} f(x) = L,$$

并以此表示, 当 x 很大的时候, $f(x)$ 变得非常接近于值 L, 并保持这种接近的状态. (更多详情请参见附录 A 的 A.3.3 节.) 重要的是要意识到, 写出 “$\lim\limits_{x\to\infty} f(x) = L$” 表示 f 的图像在 $y = L$ 处有一条右侧水平渐近线. 类似地, 当 x 趋于 $-\infty$ 时, 我们写出

$$\lim_{x\to-\infty} f(x) = L,$$

它表示当 x 变得越来越负 (或者更确切地说, $-x$ 变得越来越大) 时, $f(x)$ 会变得非常接近于值 L, 并保持接近的状态. 当然, 这对应于函数 $y = f(x)$ 的图像有一条左侧水平渐近线. 如果愿意, 你也可以把这些转化为定义:

“f 在 $y = L$ 处有一条右侧水平渐近线” 意味着 $\lim\limits_{x\to\infty} f(x) = L$.
“f 在 $y = M$ 处有一条左侧水平渐近线” 意味着 $\lim\limits_{x\to-\infty} f(x) = M$.

当然, 像 $y = x^2$ 这样的函数没有任何水平渐近线, 因为当 x 变得越来越大时, y 值只会无限上升. 用符号表示, 我们可以写作 $\lim\limits_{x\to\infty} x^2 = \infty$. 另外, 极限也有可能不存在. 例如, $\lim\limits_{x\to\infty} \sin(x)$. $\sin(x)$ 会变得越来越接近何值 (并保持这种接近状态) 呢? 它只是在 -1 和 1 之间来回振荡, 因此绝不会真正地接近任何地方. 此函数没有水平渐近线, 也不会趋于 ∞ 或 $-\infty$; 你所能作的最好回答是, $\lim\limits_{x\to\infty} \sin(x)$ 不存在 (DNE). 证明请参见附录 A 的 A.3.4 节.

让我们回到上一节看到的函数 f, 其定义为 $f(x) = \sin(1/x)$. 当 x 变得非常大时会怎么样呢? 首先, 当 x 很大时, $1/x$ 会非常接近于 0. 由于 $\sin(0) = 0$, 那么 $\sin(1/x)$ 就会非常接近于 0. x 越大, $\sin(1/x)$ 就会越来越接近于 0. 我的论证有点粗略, 但希望能说服你相信[①]

$$\lim_{x\to\infty} \sin(1/x) = 0.$$

因此, $\sin(1/x)$ 在 $y = 0$ 处有一条水平渐近线. 这就能够扩展我们之前画的 $y = \sin(1/x)$ 的图像, 至少是向右边做扩展. 我们仍旧担心当 $x < 0$ 时会发生什么. 事情不是太糟糕, 因为 f 是一个奇函数. 理由是

$$f(-x) = \sin\left(\frac{1}{-x}\right) = \sin\left(-\frac{1}{x}\right) = -\sin\left(\frac{1}{x}\right) = -f(x).$$

注意到我们使用了 $\sin(x)$ 是 x 的奇函数的事实来由 $\sin(-1/x)$ 得到 $-\sin(1/x)$. 这样一来, 由于奇函数有一个很好的性质, 就是其图像关于原点对称 (参见 1.4 节), 可以完整地画出 $y = \sin(1/x)$ 的图像, 如图 3-8 所示.

① 如果你不信, 请参见附录 A 的 A.4.1 节!

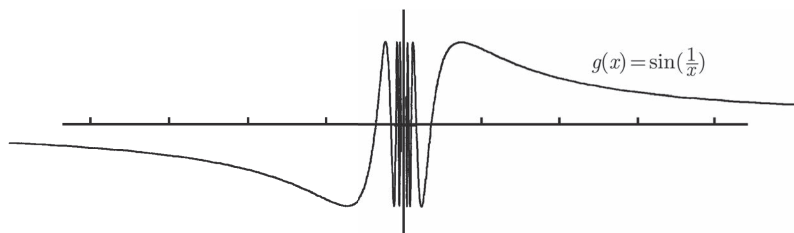


图　3-8

同样, 很难画出当 x 在 0 附近时的情况. x 越接近 0, 此函数就会振荡得越激烈. 当然, 该函数在 $x=0$ 处无意义. 在上图中, 我选择避免在中间画得密密麻麻, 而是把那里的激烈振荡留给你想象.

大的数和小的数

希望我们都认同 1 000 000 000 000 是一个大的数. 那么 −1 000 000 000 000 呢? 或许这会引起争议, 但我要让你把它看作是一个大的负数, 而不是一个小的数. 举个小的数的例子, 0.000 000 001, 同时 −0.000 000 001 也是一个小的数 (更确切地说, 是一个小的负数). 有趣的是, 我们不打算把 0 看作是个小的数：它就是零. 因此, 下面就是我们对于大的数和小的数的非正式定义：

- 如果一个数的绝对值是非常大的数, 则这个数是**大的**;
- 如果一个数非常接近于 0 (但不是真的等于 0), 则这个数是**小的**.

尽管上述定义在我们的实际应用中很有帮助, 但这实在是一个没有说服力的定义. “非常大” 和 “非常接近于 0” 分别意味着什么? 好吧, 我们考虑极限

$$\lim_{x\to\infty} f(x)=L.$$

正如之前看到的, 它表示当 x 是一个足够大的数时, $f(x)$ 的值就会几乎等于 L. 可问题是, 多大才是 “足够大” 呢? 这取决于你想让 $f(x)$ 距离 L 有多近! 不过, 从实际应用的角度出发, 如果 $y=f(x)$ 的图像看上去开始变得靠近在 $y=L$ 的水平渐近线, 那么这个数 x 足够大. 当然, 一切都依赖于函数 f 的定义, 例如图 3-9 中的两种情况.

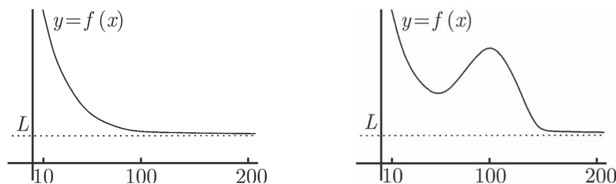


图　3-9

在这两种情况下, $f(10)$ 都不在 L 的附近. 在左图中, 当 x 至少是 100 时, $f(x)$ 看上去非常接近于 L, 因此, 任何比 100 大的数都是大数. 在右图中, $f(100)$ 远离 L,

因此, 现在的 100 就不是足够大了. 在这种情形下, 你可能需要走到 200. 那么你能够只选取一个像 1 000 000 000 000 这样的数, 然后说它已经很大了吗? 不可以, 因为一个函数有可能一直起伏不定, 直到比如 5 000 000 000 000 才变得趋于它的水平渐近线. 这里的要点是, "大的" 一词必须考虑到相关的某个函数或极限才有意义. 幸好, 没有最大, 只有更大, 往上还大有余地 —— 甚至一个像 1 000 000 000 000 这样的数, 相对于 10^{100} (古戈尔) 来说还是相当小, 而 10^{100} 与 $10^{1\,000\,000}$ 比起来又是那么微不足道 …… 顺便说一下, 我们会经常使用术语 "在 ∞ 附近" 来代替 "大的正的数". (在字面意义上说, 一个数不可能真的在 ∞ 附近, 因为 ∞ 无穷远. 不过在 $x \to \infty$ 时的极限的语境中, "在 ∞ 附近" 的说法还是说得通的.)

当然, 所有这些也都适用于 $x \to -\infty$ 时的极限, 你只需在上述所有大的正的数之前添加一个负号. 在这种情况下, 我们有时会说 "在 $-\infty$ 附近" 来强调我们所指的是大的负的数.

另一方面, 我们会经常看到极限

$$\lim_{x\to 0} f(x) = L, \quad \lim_{x\to 0^+} f(x) = L \quad 或 \quad \lim_{x\to 0^-} f(x) = L.$$

在上述三种情况下, 我们知道, 当 x 足够接近于 0 时, $f(x)$ 的值几乎是 L. (对于右极限, x 还必须为正; 而对于左极限, x 还必须为负.) 那么 x 必须离 0 多近呢? 这取决于函数 f. 因此, 当说一个数是 "小的"(或者 "接近于 0") 时, 必须结合某个函数或极限的语境来考虑, 就像在 "大的" 情形中一样.

尽管这一番讨论让之前的非正式定义确实变得更严谨了一些, 但它仍不算完美. 如果你想了解更多, 真的应该查看一下附录 A 的 A.1 节和 A.3.3 节.

3.5 关于渐近线的两个常见误解

现在是时候来纠正一些关于水平渐近线的常见误解了. 首先, 一个函数不一定要在左右两边有相同的水平渐近线. 在 3.3 节 $f(x) = 1/x$ 的图像中, 左右两侧都有 $y = 0$ 这条水平渐近线. 也就是说,

$$\lim_{x\to\infty} \frac{1}{x} = 0 \quad 和 \quad \lim_{x\to-\infty} \frac{1}{x} = 0.$$

然而, 考虑图 3-10 中 $y = \tan^{-1}(x)$ (或反三角函数 $y = \arctan(x)$, 你可以使用这两种写法中的任意一种) 的图像.

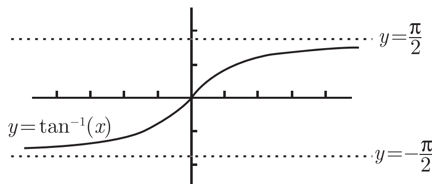


图 3-10

此函数在 $y = \pi/2$ 处有一条右侧水平渐近线, 在 $y = -\pi/2$ 处有一条左侧水平渐近线, 它们是不同的. 也可以用极限来表示：

$$\boxed{\lim_{x\to\infty} \tan^{-1}(x) = \frac{\pi}{2}} \quad \text{和} \quad \boxed{\lim_{x\to-\infty} \tan^{-1}(x) = -\frac{\pi}{2}}$$

因此, 一个函数的确可以有不同的右侧和左侧水平渐近线, 但最多只能有两条水平渐近线 (一条在右侧, 另一条在左侧). 它也有可能一条都没有, 或者只有一条. 例如, $y = 2^x$ 有一条左侧水平渐近线, 但没有右侧水平渐近线 (参见 1.6 节的图像). 这和垂直渐近线相反：一个函数可以有很多条垂直渐近线 (例如, $y = \tan(x)$ 有无穷多条垂直渐近线).

另外一个常见误解是, 一个函数不可能和它的渐近线相交. 或许你曾学到, 渐近线是一条函数越来越接近但永远不会相交的直线. 这并不正确, 至少当你讨论的是水平渐近线时. 例如, 考虑定义为 $f(x) = \sin(x)/x$ 的函数 f, 这里我们只关心当 x 是很大的正数时的函数行为. $\sin(x)$ 的值在 -1 和 1 之间振荡, 因此, $\sin(x)/x$ 的值在曲线 $y = -1/x$ 和 $y = 1/x$ 之间振荡. 此外, $\sin(x)/x$ 和 $\sin(x)$ 有相同的零点, 即 $\pi, 2\pi, 3\pi, \cdots$. 综合所有的信息, 其图像如图 3-11 所示.

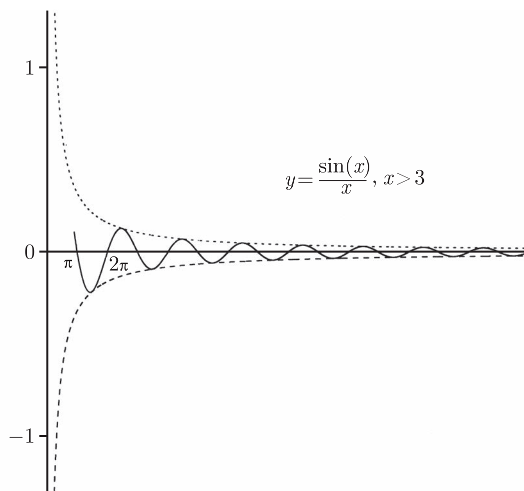


图　3-11

图像中用虚线表示的曲线 $y = 1/x$ 和 $y = -1/x$ 形成了正弦波的包络. 毫无疑问, 正如你从图像中看到的,

$$\lim_{x\to\infty} \frac{\sin(x)}{x} = 0$$

必定成立. 这意味着, x 轴是 f 的水平渐近线, 尽管 $y = f(x)$ 的图像与 x 轴一次又一次地相交. 为了证明上述极限, 我们需要应用所谓的三明治定理. 这个证明就在

下一节的结尾部分.

3.6 三明治定理



三明治定理 (又称作**夹逼定理**) 说的是, 如果一个函数 f 被夹在函数 g 和 h 之间, 当 $x \to a$ 时, 这两个函数 g 和 h 都收敛于同一个极限 L, 那么当 $x \to a$ 时, f 也收敛于极限 L.

以下是对该定理的一个更精确的描述. 假设对于所有的在 a 附近的 x, 我们都有 $g(x) \leqslant f(x) \leqslant h(x)$, 即 $f(x)$ 被夹在 $g(x)$ 和 $h(x)$ 之间. 此外, 我们假设 $\lim\limits_{x\to a} g(x) = L$ 且 $\lim\limits_{x\to a} h(x) = L$. 那么我们可以得出结论: $\lim\limits_{x\to a} f(x) = L$; 即当 $x \to a$ 时, 所有三个函数都有相同的极限. 一如往常, 一图胜千言 (见图 3-12).

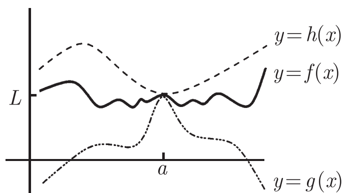


图 3-12

在图像中用实线表示的函数 f 被夹在其他两个函数 g 和 h 之间; 当 $x \to a$ 时, $f(x)$ 的极限被迫趋于 L. (三明治定理的证明参见附录 A 的 A.2.4 节.)



对于单侧极限, 我们也有一个类似版本的三明治定理, 只是这时不等式 $g(x) \leqslant f(x) \leqslant h(x)$ 仅在 a 的我们关心的一侧成立. 例如,

$$\lim_{x\to 0^+} x\sin\left(\frac{1}{x}\right)$$

是什么呢? $y = x\sin(1/x)$ 的图像和 $y = \sin(1/x)$ 的图像很相似, 只是现在, 前面有一个 x 致使函数陷于包络 $y=x$ 和 $y=-x$ 之间. 图 3-13 是 x 在 0 和 0.3 之间时的函数图像.

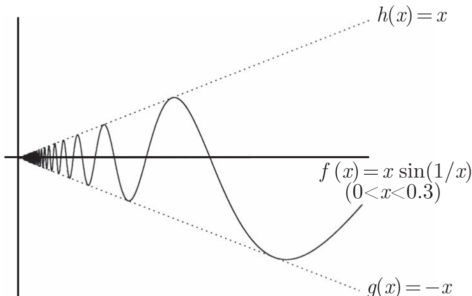


图 3-13

从图中可以看到, 当 x 趋于 0 时, 函数仍旧有激烈的振荡, 但现在它们被包络线抑制

着. 特别是, 这里求我们想要的极限正是三明治定理的一个完美应用. 函数 g 是下方的包络线 $y=-x$, 而函数 h 是上方的包络线 $y=x$. 我们需要证明对于 $x>0$, 有 $g(x)\leqslant f(x)\leqslant h(x)$. 由于只需要 $f(x)$ 在 $x=0$ 处的右极限, 所以我们不关心 $x<0$ 时的情况. (事实上, 如果扩展到 x 轴负半轴, 你可以看到, 对于 $x<0$, $g(x)$ 实际大于 $h(x)$, 所以三明治要翻个身!) 那么当 $x>0$ 时, 要怎样证明 $g(x)\leqslant f(x)\leqslant h(x)$ 呢? 我们将会用到任意数 (在我们的例子中是 $1/x$) 的正弦都处于 -1 和 1 之间这一事实:

$$-1\leqslant \sin\left(\frac{1}{x}\right)\leqslant 1.$$

现在用 x 乘以这个不等式, 由于 $x>0$, 得到

$$-x\leqslant x\sin\left(\frac{1}{x}\right)\leqslant x.$$

而这正是我们需要的 $g(x)\leqslant f(x)\leqslant h(x)$. 最后, 注意到

$$\lim_{x\to 0^+} g(x)=\lim_{x\to 0^+}(-x)=0 \quad 及 \quad \lim_{x\to 0^+} h(x)=\lim_{x\to 0^+} x=0.$$

因此, 由于当 $x\to 0^+$ 时, 夹逼的函数 $g(x)$ 和 $h(x)$ 的值收敛于同一个数 0, 所以 $f(x)$ 也一样. 也就是说, 证明了

$$\lim_{x\to 0^+} x\sin\left(\frac{1}{x}\right)=0.$$

要记住, 如果前面没有因子x, 上式显然不成立; 正如我们在3.3节看到的, 当 $x\to 0^+$ 时, $\sin(1/x)$ 的极限不存在.

我们还没有解决上一节结尾部分的那个极限证明问题! 回想一下, 要证明的是

$$\lim_{x\to\infty}\frac{\sin(x)}{x}=0.$$

为了证明此式, 需要用到三明治定理一个稍有不同的形式, 涉及在 ∞ 处的极限. 在这种情况下, 如果对于所有的很大的 x, 都有 $g(x)\leqslant f(x)\leqslant h(x)$ 成立; 又如果已知 $\lim\limits_{x\to\infty} g(x)=L$ 且 $\lim\limits_{x\to\infty} h(x)=L$. 就可以说, $\lim\limits_{x\to\infty} f(x)=L$. 这与有限处极限的三明治定理几乎是一样的. 为了确立上述极限, 还要用到, 对于所有的 x, 都有 $-1\leqslant\sin(x)\leqslant 1$, 但这次, 对于所有的 $x>0$, 要用该不等式除以 x 得到

$$-\frac{1}{x}\leqslant\frac{\sin(x)}{x}\leqslant\frac{1}{x}.$$

现在, 令 $x\to\infty$, 由于 $-1/x$ 和 $1/x$ 的极限都是 0, $\sin(x)/x$ 的极限也必为 0. 也就是说, 由于

$$\lim_{x\to\infty}-\frac{1}{x}=0 \quad 和 \quad \lim_{x\to\infty}\frac{1}{x}=0,$$

也必有

$$\lim_{x\to\infty}\frac{\sin(x)}{x}=0.$$

综上, 三明治定理说的是：

> 如果对于所有在 a 附近的 x 都有 $g(x) \leqslant f(x) \leqslant h(x)$, 且 $\lim\limits_{x\to a} g(x) = \lim\limits_{x\to a} h(x) = L$, 则 $\lim\limits_{x\to a} f(x) = L$.

这也适用于左极限或右极限; 在那种情况下, 不等式只需要在 a 的相应一侧对于 x 成立即可. 当 a 是 ∞ 或 $-\infty$ 时它也适用; 在那种情况下, 要求对于所有的非常大的 (分别是正的或负的)x, 不等式成立.

3.7 极限的基本类型小结

我们已经看过了极限的多种基本类型. 下面展示一些各种基本类型的代表性图像, 以此来结束本章.



(1) 在 $x = a$ 时的右极限, 见图 3-14. 这时在 $x = a$ 的左侧以及 $x = a$ 处 $f(x)$ 的行为是无关紧要的. (也就是说, 当讨论右极限时, 对于 $x \leqslant a$, $f(x)$ 取何值都不要紧. 事实上, 对于 $x \leqslant a$, $f(x)$ 甚至不需要被定义.)

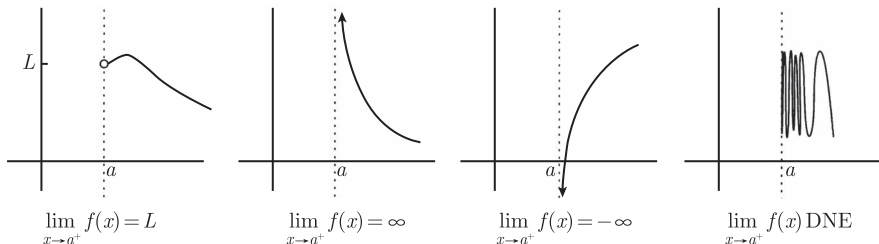


图 3-14

(2) 在 $x = a$ 时的左极限, 见图 3-15. 这时在 $x = a$ 的右侧以及 $x = a$ 处 $f(x)$ 的行为是无关紧要的.

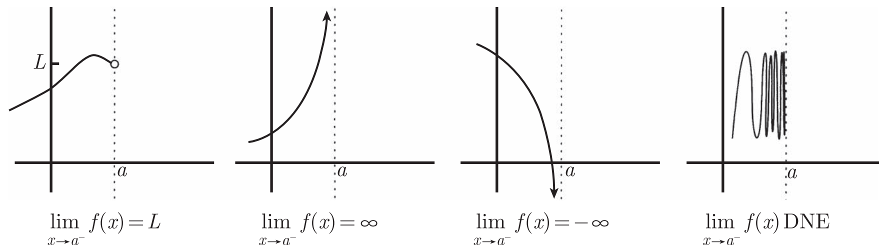


图 3-15

(3) 在 $x = a$ 时的双侧极限, 见图 3-16. 在左图中, 左极限和右极限存在但不相等, 因此, 双侧极限不存在. 在右图中, 左极限和右极限存在并相等, 因此, 双侧极限存在并等于左右极限值. $f(a)$ 的值是无关紧要的.

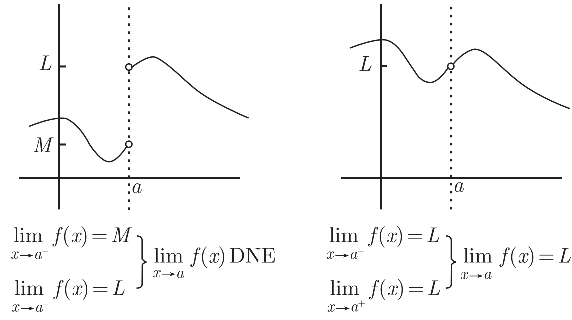


图 3-16

(4) 在 $x \to \infty$ 时的极限, 见图 3-17.

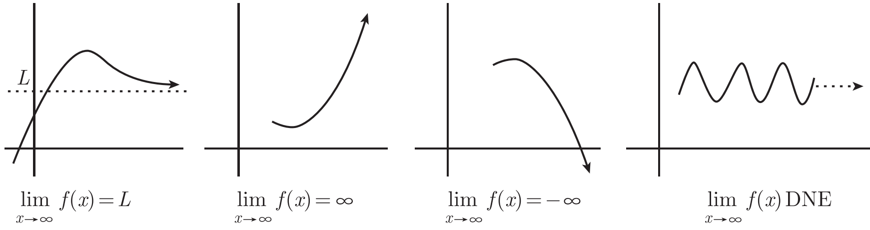


图 3-17

(5) 在 $x \to -\infty$ 时的极限, 见图 3-18.

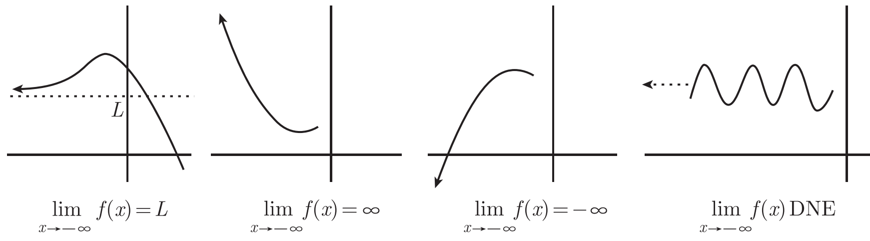


图 3-18

第 4 章　求解多项式的极限问题

在上一章中, 我们主要是从概念的角度学习了极限. 现在是时候来看一看求解极限的一些技巧了. 目前, 我们将注意力集中在涉及多项式的极限问题上; 以后, 我们还会看到如何处理涉及三角函数、指数函数和对数函数的极限问题. 正如我们将在下一章看到的, 微分会涉及比率的极限, 因此, 我们的注意力将主要集中在这种类型的极限上.

当你取两个多项式的比的极限时, 真正要紧的是注意到极限是在哪里取的. 特别是, 处理 $x \to \infty$ 和处理 $x \to a$(对于某个有限的数 a) 的技巧是完全不同的. 因此, 我们将对涉及下列函数类型的极限分开研究:

- $x \to a$ 时的有理函数;
- $x \to a$ 时的涉及平方根的函数;
- $x \to \infty$ 时的有理函数;
- $x \to \infty$ 时的类多项式 (或 “多项式型”) 函数的比;
- $x \to -\infty$ 时的有理函数/多项式型函数;
- 涉及绝对值的函数.

4.1　$x \to a$ 时的有理函数的极限

让我们以极限

$$\lim_{x \to a} \frac{p(x)}{q(x)}$$

开始吧, 其中 p 和 q 都是多项式, 并且 a 是一个有限的数. (记住, 两个多项式之比 $p(x)/q(x)$ 被称作有理函数.) 你首先总是应该尝试用 a 的值替换 x. 如果分母不为 0, 那么你一切顺利, 极限值就是你做替换后所得到的值. 例如, 极限



$$\lim_{x \to -1} \frac{x^2 - 3x + 2}{x - 2}$$

是什么呢? 可以简单地将 $x = -1$ 代入表达式 $\left(x^2 - 3x + 2\right)/(x - 2)$ 中, 得到

$$\frac{(-1)^2 - 3(-1) + 2}{-1 - 2} = \frac{6}{-3} = -2.$$

其分母不为 0, 因此, -2 就是极限值. (我知道我在上一章说过, 函数在极限点上的值, 在上述情况下, 就是在 $x = -1$ 处的值, 是无关紧要的; 但在下一章中, 我们将学到连续性的概念, 它将证明这种 “代入” 法是没有问题的.)



另一方面, 如果你想要求

$$\lim_{x\to 2}\frac{x^2-3x+2}{x-2},$$

那么代入 $x=2$ 并不会起到很好的效果：你会得到 $(4-6+2)/(2-2)$, 简化为 $0/0$. 这被称作**不定式**. 如果你使用代入法并得到零比零的形式, 那么什么都可能会发生：极限或许是有限的, 极限或许是 ∞ 或 $-\infty$, 或者极限或许不存在. 我们可以借助因式分解这一重要技巧来求解上例. 特别是, x^2-3x+2 可以被分解为 $(x-2)(x-1)$. 因此, 通过删除公因子我们可以写出

$$\lim_{x\to 2}\frac{x^2-3x+2}{x-2}=\lim_{x\to 2}\frac{(x-2)(x-1)}{x-2}=\lim_{x\to 2}(x-1).$$

现在, 就可以将 $x=2$ 代入到表达式 $(x-1)$ 中了; 你会得到 $2-1$, 其结果是 1. 那就是要求的极限值.



这引出了经常会被误解的一点. 有两个函数 f 和 g, 定义分别是

$$f(x)=\frac{x^2-3x+2}{x-2}\quad 和 \quad g(x)=x-1.$$

那它们是同一个函数吗？为什么不能说

$$f(x)=\frac{x^2-3x+2}{x-2}=\frac{(x-2)(x-1)}{x-2}=x-1=g(x)?$$

好吧, 你几乎可以这么说! 唯一的问题出在当 $x=2$ 时, 因为那时分母 $(x-2)$ 就等于 0, 而这就说不通了. 因此, f 和 g 不是同一个函数：数 2 不在 f 的定义域中, 但它却在 g 的定义域中. (事实上, 之前已经碰到过这个函数 f, 可参见第 3 章开头的讨论及图像.) 另一方面, 如果你把极限符号放在以上等式链中每一项的最前面, 那么一切就都没问题, 因为这时, $f(x)$ 和 $g(x)$ 在 $x=2$ 处的值是无关紧要的, 只有那些在 $x=2$ 附近的 $f(x)$ 和 $g(x)$ 的值才有关紧要. 因此, 上述极限问题的解的确是有效的.

来看看另一个有关不定式的例子. 同样, 这里的技巧是试着将所有多项式做因式分解. 为此, 除了要知道如何分解二次多项式之外, 了解立方差的公式也非常重要：

$$\boxed{a^3-b^3=(a-b)(a^2+ab+b^2).}$$



以下是一个更难的例子, 你需要使用上述公式. 求

$$\lim_{x\to 3}\frac{x^3-27}{x^4-5x^3+6x^2}.$$

如果你将 $x=3$ 代入, 你会得到 $0/0$(试着做一下就会知道了). 因此让我们试着来分解分子和分母. 分子是 x^3 和 3^3 的差, 因此, 可以使用上述的加框公式. 分母有一个明显的因子是 x^2, 因此它可以被写成 $x^2\left(x^2-5x+6\right)$. 二次的 x^2-5x+6 也

可以被分解; 综上, 你可以验证一下, 有

$$\lim_{x \to 3} \frac{x^3 - 27}{x^4 - 5x^3 + 6x^2} = \lim_{x \to 3} \frac{(x-3)(x^2+3x+9)}{x^2(x-3)(x-2)}.$$

代入 $x = 3$ 不起作用, 因为因子 $(x-3)$ 在分母上. 另一方面, 由于是要取极限, 只需要关注 x 在 3 附近的情况; 因此, 能够消去分子和分母中的公因子 $(x-3)$ (它们永远不会等于 0). 因此, 在因式分解并消去公因子之后使用代入法, 完整的求解为

$$\lim_{x \to 3} \frac{x^3 - 27}{x^4 - 5x^3 + 6x^2} = \lim_{x \to 3} \frac{(x-3)(x^2+3x+9)}{x^2(x-3)(x-2)} = \lim_{x \to 3} \frac{x^2+3x+9}{x^2(x-2)}$$
$$= \frac{3^2 + 3 \cdot 3 + 9}{3^2(3-2)} = 3.$$

要是分母为 0 但分子不为 0 又会怎么样呢? 在那种情况下, 将总会牵扯到一条垂直渐近线, 即有理函数的图像在你感兴趣的 x 值上会有一条垂直渐近线. 但这里的问题是, 会有四种情形出现. 在图 4-1 所示的每一幅图里, f 是一个我们关心的有理函数, 图下面则是 $x = a$ 处的各种极限.

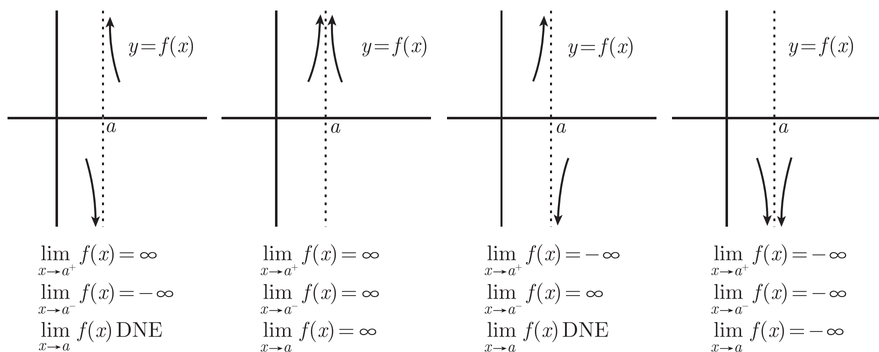


图 4-1

那么你又如何分辨出你在处理的是这四种情形中的哪一种呢? 其实, 你只需要查看一下 $f(x)$ 在 $x = a$ 两边的符号就可以了. 例如, 如果它在两边都是正的, 那么你一定是在处理上述的第二种情形. 下面就是一个实际的例子: 如何求



$$\lim_{x \to 1} \frac{2x^2 - x - 6}{x(x-1)^3}?$$

首先, 代入 $x = 1$ 得出 $-5/0$. (自己尝试做一下!) 因此, 我们必定是在处理上述四种情形中的一种. 会是哪一种呢? 我们指定 $f(x) = \left(2x^2 - x - 6\right) / \left(x(x-1)^3\right)$, 并观察当移动 x 到 1 的附近时会有什么情况发生. 首先注意到的是, 当 $x = 1$ 时, 分子 $\left(2x^2 - x - 6\right)$ 等于 -5, 因此, 当在 1 的附近稍微移动一下 x, 则分子保持负值. 那么分母里的因子 x 会怎样呢? 当 $x = 1$ 时, 这个因子当然是 1, 它是正的. 并且, 当你在 1 的附近稍微移动一下 x, 它也保持为正的. 关键因子是 $(x-1)^3$, 当 $x > 1$

时它为正, 而当 $x<1$ 时为负. 因此, 可以总结如下 (使用 $(+)$ 和 $(-)$ 分别表示正的和负的量, 并且当然要利用 $(-)\cdot(-)=(+)$ 等事实):

$$当\ x>1, \frac{(-)}{(+)\cdot(+)}=(-); \quad 当\ x<1, \frac{(-)}{(+)\cdot(-)}=(+).$$

也就是说, 当 x 比 1 大一点的时候, $f(x)$ 是负的; 而当 x 比 1 小一点的时候, $f(x)$ 是正的. 比对上述的四幅图, 只有第三幅图对应我们的问题. 特别是, 我们可以看到双侧极限

$$\lim_{x\to1}\frac{2x^2-x-6}{x(x-1)^3}$$

不存在, 而单侧极限存在 (尽管它们是无穷大); 具体来说,

$$\lim_{x\to1^+}\frac{2x^2-x-6}{x(x-1)^3}=-\infty \quad 和 \quad \lim_{x\to1^-}\frac{2x^2-x-6}{x(x-1)^3}=\infty.$$

现在, 假设对极限做了微小的改变, 使它成为

$$\lim_{x\to1}\frac{2x^2-x-6}{x(x-1)^2}.$$

一切又会怎样呢? 当 x 接近于 1 时, 分子仍然是负的, 且因子 x 依然是正的, 但 $(x-1)^2$ 呢? 由于它是一个平方, 当 x 接近 1 但不等于 1 时, 它必定是正的. 因此, 现在有下列情形:

$$当\ x>1, \frac{(-)}{(+)\cdot(+)}=(-); \quad 当\ x<1, \frac{(-)}{(+)\cdot(+)}=(-).$$

现在在 $x=1$ 的两边有负值, 因此必定有

$$\lim_{x\to1}\frac{2x^2-x-6}{x(x-1)^2}=-\infty.$$

当然, 左极限和右极限也都是 $-\infty$.

4.2 $x\to a$ 时的平方根的极限

考虑极限

$$\lim_{x\to5}\frac{\sqrt{x^2-9}-4}{x-5}.$$

如果代入 $x=5$, 你会得到 0/0 型的不定式 (试着做一下看看!). 进行因式分解也好像不太管用 —— 你可以将 x^2-9 写作 $(x-3)(x+3)$, 但这也不会起多大作用, 因为还有一个 -4 在分子上. 你需要做的是, 把分子分母同时乘以 $\sqrt{x^2-9}+4$, 也就是 $\sqrt{x^2-9}-4$ 的**共轭表达式**. (或许你已经在之前的数学学习过程中碰到过共轭表达式了, 尤其是在分母有理化的时候. 其基本思想是, $a-b$ 的共轭表达式是 $a+b$, 反之亦然.) 因此, 得到

$$\lim_{x\to5}\frac{\sqrt{x^2-9}-4}{x-5}=\lim_{x\to5}\frac{\sqrt{x^2-9}-4}{x-5}\times\frac{\sqrt{x^2-9}+4}{\sqrt{x^2-9}+4}.$$

这看起来更复杂了, 但某种好事情即将发生：使用公式 $(a-b)(a+b)=a^2-b^2$, 分子可简化为 $\left(\sqrt{x^2-9}\right)^2-4^2$, 即 x^2-25. 因此, 以上极限就是

$$\lim_{x\to 5}\frac{x^2-25}{(x-5)(\sqrt{x^2-9}+4)}.$$

将 x^2-25 分解为 $(x-5)(x+5)$ 并消去分子分母中的公因子, 此极限变为

$$\lim_{x\to 5}\frac{(x-5)(x+5)}{(x-5)(\sqrt{x^2-9}+4)}=\lim_{x\to 5}\frac{x+5}{\sqrt{x^2-9}+4}.$$

现在, 如果代入 $x=5$ 就没有问题了, 你会得到 10/8, 即 5/4. 这个例子的要义在于, 如果你碰到一个平方根加上或减去另外一个量, 可以试着把分子分母同时乘以其共轭表达式, 也许会有令人高兴的惊喜发生呢!

4.3 $x \to \infty$ 时的有理函数的极限

现在, 让我们回到有理函数, 但这次要看看当 $x\to\infty$ 而不是某个有限的值时会有什么情况发生. 用符号表示, 现在想要求极限

$$\lim_{x\to\infty}\frac{p(x)}{q(x)},$$

其中 p 和 q 是多项式. 现在, 这里有一个非常重要的多项式性质：当 x 很大时, 首项决定一切. 这就是说, 如果你有一个多项式 p, 那么当 x 变得越来越大时, $p(x)$ 的表现就好像只有它的首项存在一样. 例如, 假设 $p(x)=3x^3-1000x^2+5x-7$. 让我们设 $p_L(x)=3x^3$, 它就是 p 的首项. 这里我要说的是：当 x 变得非常非常大时, $p(x)$ 和 $p_L(x)$ 会相对地非常接近. 更确切地说, 我们有

$$\lim_{x\to\infty}\frac{p(x)}{p_L(x)}=1.$$

在明白为什么上式成立之前, 先来看一下它想要表达的意义. 想象一下如果没有极限符号, 这个等式将是

$$\frac{p(x)}{p_L(x)}=1,$$

这意味着 $p(x)=p_L(x)$. 很明显这不是真的 (至少对于绝大多数的 x 值来说), 但随着 x 越来越大, 该等式就会越来越趋近于是真的. 那么为什么不写成

$$\lim_{x\to\infty}p(x)=\lim_{x\to\infty}p_L(x)?$$

这确实是真的, 但由于两边都是 ∞, 它毫无意义. 因此, 我们只能接受, 用其比接近于 1 来表达 $p(x)$ 和 $p_L(x)$ 非常接近. 随着 x 越来越大, 其比趋于 1, 而不必等于 1.



这说得通吗？为什么是首项呢？为什么不是其他项中的一项呢？如果你着急, 可以跳到下一段去看看数学证明; 然而, 首先, 我想让你有个直观感受, 用实际的大的 x 值做检验, 来看看在例子 $p(x)=3x^3-1000x^2+5x-7$ 中会发生什么. 从 $x=100$ 开始. 在那种情况下, $3x^3$ 是三百万, 而 $1000x^2$ 是一千万. 量 $5x$ 仅为 500, 而 7 更

是无足轻重. 因此, 所有的值加在一起, 可以看到 $p(100)$ 大概是负七百万. 另一方面, $p_L(100)$ 是三百万, 这看起来不是太好: $p(100)$ 和 $p_L(100)$ 完全不同. 可是不要丧失信心, 毕竟 100 不是很大. 假设我们将 x 设为 1 000 000, 即一百万. 那么 $3x^3$ 会变得非常大: 它是 3 000 000 000 000 000 000, 即三百万万亿! 相比之下, $1000x^2$ 会变得相对微小, 它仅是一千万亿 (即 1 000 000 000 000 000), 而 $5x$ 只是五百万, 更是微不足道. 项 -7 就只是让人可发一笑了. 因此, 为了计算 $p(1\,000\,000)$, 需要用三百万万亿减去一千万亿再加上一些微小的变化 (在五百万以下的小变化). 不得不承认, 结果还是接近于三百万万亿! 毕竟, 这里处理的是多少个万亿呢? 我们有三百万个万亿, 而只是去掉了其中的一千个, 所以剩下还有差不多三百万个万亿. 也就是说, $p(1\,000\,000)$ 大概是三百万万亿, 这不就是 $p_L(1\,000\,000)$ 的值吗? 这里的要点是, 当 x 变大时, 最高次数项比其他项增长得更快. 事实上, 如果你用一个更大的数来代替 1 000 000, x^3 与诸如 x^2 和 x 这样的低次数项之间的差异会变得更为明显.

书归正传, 让我们试着给出一个真正的证明, 证明

$$\lim_{x\to\infty}\frac{p(x)}{p_L(x)}=1.$$

这里必须要做一些实际的数学了. 先写出

$$\lim_{x\to\infty}\frac{p(x)}{p_L(x)}=\lim_{x\to\infty}\frac{3x^3-1000x^2+5x-7}{3x^3},$$

它可简化为

$$\lim_{x\to\infty}\left(\frac{3x^3}{3x^3}-\frac{1000x^2}{3x^3}+\frac{5x}{3x^3}-\frac{7}{3x^3}\right)=\lim_{x\to\infty}\left(1-\frac{1000}{3x}+\frac{5}{3x^2}-\frac{7}{3x^3}\right).$$

你该如何处理它呢? 首先注意到, 可以将最后一个表达式分成四个单独的极限. 因此, 如果你知道, 当 x 变得非常大时, 1, $-1000/3x$, $5/3x^2$ 和 $-7/3x^3$ 这四个量会发生什么情况的话, 那么就可以把这四个极限加在一起来得到你想要求的极限. 技术上讲, 这可以描述为 "和的极限等于极限的和"; 这在所有的极限都是有限的[①]时候成立. 因此, 我们要分别考虑这四个量. 第一个是 1, 不管 x 是什么, 它总是 1. 第二个量是 $-1000/3x$. 当 x 变大时, 它会怎么样呢? 也就是说,

$$\lim_{x\to\infty}-\frac{1000}{3x}$$

是什么呢? 这里的诀窍是, 意识到你可以将因子 $-1000/3$ 提出来. 特别是, 该极限可以表示为

$$\lim_{x\to\infty}-\frac{1000}{3}\frac{1}{x}.$$

① 如果极限不是有限的, 它就不成立! 试考虑 $\lim\limits_{x\to\infty}(x+(1-x))$. 对于任意的 x, 都有 $(x+(1-x))=1$, 因此, 此极限是 1. 另一方面, 这两个单独的 (x) 和 $(1-x)$ 的极限是 $\lim\limits_{x\to\infty}(x)$ 和 $\lim\limits_{x\to\infty}(1-x)$. 第一个极限是 ∞, 第二个极限是 $-\infty$, 但 $\infty+(-\infty)=1$ 不成立. 事实上, 表达式 $\infty+(-\infty)$ 是无意义的.

由于 $-1000/3$ 是常数, 不管 x 是什么, 它都不会改变. 因此, 把它拖到极限符号之外 (更多详情参见附录 A 的 A.2.2 节). 于是有

$$\lim_{x\to\infty} -\frac{1000}{3}\frac{1}{x} = -\frac{1000}{3} \lim_{x\to\infty} \frac{1}{x}.$$

我们已经知道, 一个非常大的数的倒数是一个非常小的数 (记住, 这意味着一个非常接近于零的数). 因此, $\lim\limits_{x\to\infty} 1/x = 0$, 而 $-1000/3$ 乘上此极限还是 0. 于是结论是

$$\lim_{x\to\infty} -\frac{1000}{3x} = 0.$$

事实上, 你应该直接写出上述结论, 而不用加入过多细节. 更一般地, 你可以使用下述定理：对于任意的 $n > 0$, 只要 C 是常数, 就有

$$\boxed{\lim_{x\to\infty} \frac{C}{x^n} = 0.}$$

由这个事实可知, 当 x 变得非常大时, 其他两项 $5/3x^2$ 和 $-7/3x^3$ 也趋于 0. 因此, 完整的论证是

$$\begin{aligned}\lim_{x\to\infty} \frac{3x^3 - 1000x^2 + 5x - 7}{3x^3} &= \lim_{x\to\infty} \left(1 - \frac{1000}{3x} + \frac{5}{3x^2} - \frac{7}{3x^3}\right)\\ &= 1 - 0 + 0 + 0 = 1.\end{aligned}$$

这样就证明了, 在特例 $p\,(x) = 3x^3 - 1000x^2 + 5x - 7$ 中,

$$\lim_{x\to\infty} \frac{p(x)}{p(x)\text{的首项}} = 1.$$

幸运的是, 同样的方法适用于任意的多项式, 并且我们会在本章的剩余部分反复使用它.

方法和例子



此方法的一般思想是：当看到某个关于 p 的多项式 $p\,(x)$ 是多于一项时, 把它代以

$$\frac{p(x)}{p(x)\text{的首项}} \times (p(x)\text{的首项}).$$

对于每一个多项式都这样做! 注意到, 所需做的就是用该多项式除以并乘以其首项, 因此并没有改变 $p\,(x)$ 的量. 这里的要点是, 当 $x \to \infty$ 时, 以上表达式中的分式的极限是 1, 并且首项比原来的表达式简单得多. 让我们来看看这在实际中如何应用吧. 例如,



$$\lim_{x\to\infty} \frac{x - 8x^4}{7x^4 + 5x^3 + 2000x^2 - 6}$$

是什么呢? 有两个多项式：一个在上, 一个在下. 对于分子, 首项是 $-8x^4$. (不要被分子中项的顺序所迷惑, 首项并不总是写在最前面!) 因此, 要把分子代以

$$\frac{x-8x^4}{-8x^4}\times(-8x^4).$$

类似地, 分母的首项是 $7x^4$, 因此, 把分母代以

$$\frac{7x^4+5x^3+2000x^2-6}{7x^4}\times(7x^4).$$

做完这两次替换, 会有

$$\lim_{x\to\infty}\frac{x-8x^4}{7x^4+5x^3+2000x^2-6}=\lim_{x\to\infty}\frac{\dfrac{x-8x^4}{-8x^4}\times(-8x^4)}{\dfrac{7x^4+5x^3+2000x^2-6}{7x^4}\times(7x^4)}.$$

对于这个式子, 你应该关注的是比

$$\frac{-8x^4}{7x^4},$$

因为这是精华所在. 其他分式的极限都是 1, 但我们已经有效地从两个多项式中 “压榨” 出了所有重要的 “果汁”, 得到简单的两个首项的比. 幸运的是, 那个比可简化为 $-8/7$, 这应该就是我们的答案了. 确凿起见, 必须证明其他分式的极限为 1, 但这不成问题. 你看, 在每一个小的分式里, 可以做除法, 并且上述极限可以写作

$$\lim_{x\to\infty}\frac{-\dfrac{1}{8x^3}+1}{1+\dfrac{5}{7x}+\dfrac{2000}{7x^2}-\dfrac{6}{7x^4}}\times\frac{-8x^4}{7x^4}.$$

现在来取极限. 根据上一节的方框公式, 当 $x\to\infty$ 时, 任何形如 C/x^n 的表达式都趋于 0(只要 C 是常数, 且 $n>0$). 因此, 大多数的项就消失了! 我们也可以消去右边的因子 x^4, 将上式简化为

$$\frac{0+1}{1+0+0-0}\times\frac{-8}{7}=\frac{1}{1}\times\frac{-8}{7}=\frac{-8}{7}.$$

这样就算完成本题了.



这里还有另外一个例子：求

$$\lim_{x\to\infty}\frac{(x^4+3x-99)(2-x^5)}{(18x^7+9x^6-3x^2-1)(x+1)}.$$



这里有四个多项式, 首项分别是 x^4, $-x^5$, $18x^7$ 及 x. 因此, 将对其中的每一个多项式来使用我们的方法! 在继续阅读之前, 试着自己做一下看看. 即使你不做, 也要确保你理解了以下论证过程中的每一步：

$$\begin{aligned}
&\lim_{x\to\infty}\frac{(x^4+3x-99)(2-x^5)}{(18x^7+9x^6-3x^2-1)(x+1)}\\
&=\lim_{x\to\infty}\frac{\left(\dfrac{x^4+3x-99}{x^4}\times(x^4)\right)\left(\dfrac{2-x^5}{-x^5}\times(-x^5)\right)}{\left(\dfrac{18x^7+9x^6-3x^2-1}{18x^7}\times(18x^7)\right)\left(\dfrac{x+1}{x}\times(x)\right)}\\
&=\lim_{x\to\infty}\frac{\left(1+\dfrac{3}{x^3}-\dfrac{99}{x^4}\right)\left(-\dfrac{2}{x^5}+1\right)}{\left(1+\dfrac{9}{18x}-\dfrac{3}{18x^5}-\dfrac{1}{18x^7}\right)\left(1+\dfrac{1}{x}\right)}\times\frac{(x^4)(-x^5)}{(18x^7)(x)}\\
&=\frac{(1+0-0)(0+1)}{(1+0-0-0)(1+0)}\times\lim_{x\to\infty}\frac{-x}{18}=\lim_{x\to\infty}\frac{-x}{18}=-\infty.
\end{aligned}$$

这里的要点是, 我们萃取出各首项, 写成比

$$\frac{(x^4)(-x^5)}{(18x^7)(x)},$$

而它可化简为 $-x/18$. 其他的都不会产生影响! 最后, 当 $x\to\infty$ 时, $-x/18$ 趋于 $-\infty$, 因此, 它就是我们想要求的极限的 “值”.

在前两个例子中, 我们看到极限有可能是有限的且非零 (得到答案 $-8/7$), 也有可能是无限的 (得到答案 $-\infty$). 现在来看一下在这些例子中的多项式的次数吧. 在第一个例子中, 分子和分母的次数都是 4. 在第二个例子中, 分子是次数为 4 和 5 的多项式的乘积, 如果把它们乘出来, 会得到一个次数为 9 的多项式. 类似地, 分母是次数为 7 和 1 的多项式的乘积, 因此, 它的总次数是 8. 在这种情况下, 分子的次数大于分母的次数. 另一方面, 试考虑极限

$$\lim_{x\to\infty}\frac{2x+3}{x^2-7}.$$

用我们的方法来求解：

$$\begin{aligned}
\lim_{x\to\infty}\frac{2x+3}{x^2-7}&=\lim_{x\to\infty}\frac{\dfrac{2x+3}{2x}\times(2x)}{\dfrac{x^2-7}{x^2}\times(x^2)}=\lim_{x\to\infty}\left(\frac{1+\dfrac{3}{2x}}{1-\dfrac{7}{x^2}}\right)\times\frac{2x}{x^2}\\
&=\frac{1+0}{1-0}\times\lim_{x\to\infty}\frac{2}{x}=0.
\end{aligned}$$

这里, 分母的次数为 2, 大于分子的次数 (为 1). 结果是, 分母占主导, 因此极限为 0. 一般地, 考虑极限

$$\lim_{x\to\infty}\frac{p(x)}{q(x)},$$

其中 p 和 q 为多项式, 我们可以说：

(1) 如果 p 的次数等于 q 的次数, 则极限是有限的且非零;

(2) 如果 p 的次数大于 q 的次数, 则极限是 ∞ 或 $-\infty$;

(3) 如果 p 的次数小于 q 的次数, 则极限是 0.

(当 $x \to -\infty$, 相应极限为

$$\lim_{x\to-\infty} \frac{p(x)}{q(x)}$$

时, 所有这些也成立, 4.5 节将考虑这种情况.) 使用我们的方法可以很容易地证明这些事实. 不过尽管这些事实很有用, 但你并不需要用它们来解题; 你应该使用前面教的乘除方法, 然后使用这些事实来检验你的答案是否说得通.

4.4 $x \to \infty$ 时的多项式型函数的极限

考虑函数 f, g 和 h, 此三个函数分别被定义为

$$f(x) = x^3 + 4x^2 - 5x^{2/3} + 1,\ g(x) = \sqrt{x^9 - 7x^2 + 2},$$
$$h(x) = x^4 - \sqrt{x^3 + \sqrt[5]{x^2 - 2x + 3}}.$$

这些都不是多项式, 因为它们含有分数次数或 n 次根, 但它们看起来有点像多项式. 事实上, 上一节的方法也适用于这类对象. 因此, 我称它们为 "多项式型函数".

处理多项式型函数的原理与处理多项式的类似, 只是这次首项是什么可能不会那么清晰. 平方根 (或立方根、四次根等) 的出现会造成很大干扰. 例如, 让我们考虑

$$\lim_{x\to\infty} \frac{\sqrt{16x^4 + 8} + 3x}{2x^2 + 6x + 1}.$$

分母是一个带有首项 $2x^2$ 的多项式, 因此, 我们可以代之以

$$\frac{2x^2 + 6x + 1}{2x^2} \times (2x^2).$$

那么分子怎么办呢? 在平方根符号下的部分是多项式 $16x^4 + 8$, 且它的首项为 $16x^4$. 如果你对其取平方根, 你会得到 $4x^2$. 因此, 你应该想象分子就像是 $4x^2 + 3x$. 它的首项为 $4x^2$, 所以我们就用它了. 具体地, 我们把分子代以

$$\frac{\sqrt{16x^4 + 8} + 3x}{4x^2} \times (4x^2).$$

你又该如何化简第一个分式呢? 答案是, 你可以把 $4x^2$ 拖进平方根符号, 它就变为 $16x^4$:

$$\frac{\sqrt{16x^4 + 8} + 3x}{4x^2} = \frac{\sqrt{16x^4 + 8}}{4x^2} + \frac{3x}{4x^2} = \sqrt{\frac{16x^4 + 8}{16x^4}} + \frac{3x}{4x^2}.$$

通过进一步拆分和消去, 可以将其化简为

$$\sqrt{1 + \frac{8}{16x^4}} + \frac{3}{4x}.$$

当 $x \to \infty$ 时, 分母中包含 x 的部分就消失了. 因此, 该表达式趋于

$$\sqrt{1 + 0} + 0 = 1.$$

最后, 将所有的放在一起, 写出原始问题的解:

$$\lim_{x\to\infty}\frac{\sqrt{16x^4+8}+3x}{2x^2+6x+1}=\lim_{x\to\infty}\frac{\dfrac{\sqrt{16x^4+8}+3x}{4x^2}\times(4x^2)}{\dfrac{2x^2+6x+1}{2x^2}\times(2x^2)}$$

$$=\lim_{x\to\infty}\frac{\sqrt{\dfrac{16x^4+8}{16x^4}}+\dfrac{3x}{4x^2}}{\dfrac{2x^2+6x+1}{2x^2}}\times\frac{4x^2}{2x^2}=\lim_{x\to\infty}\frac{\sqrt{1+\dfrac{8}{16x^4}}+\dfrac{3}{4x}}{1+\dfrac{6}{2x}+\dfrac{1}{2x^2}}\times\frac{4}{2}$$

$$=\frac{\sqrt{1+0}+0}{1+0+0}\times 2=2.$$

这很棒, 不是吗? 看上去很乱, 但确实很棒. 现在, 来看看当将情形稍加修改后会发生什么. 试考虑

$$\lim_{x\to\infty}\frac{\sqrt{16x^4+8}+3x^3}{2x^2+6x+1}.$$

唯一的变化是, 上例分子中的项 $3x$ 变成了 $3x^3$. 这会有什么影响呢? 好吧, 我们曾说过, 对于很大的 x, $\sqrt{16x^4+8}$ 这一项就像是 $4x^2$. 但这一次, 更高次数的项 $3x^3$ 超过了它. 因此, 现在必须把分子代以

$$\frac{\sqrt{16x^4+8}+3x^3}{3x^3}\times(3x^3);$$

当然, 当把 $3x^3$ 拖进平方根符号时, 它会变为 $9x^6$. 将所有的放在一起, 得到问题的解如下:

$$\lim_{x\to\infty}\frac{\sqrt{16x^4+8}+3x^3}{2x^2+6x+1}=\lim_{x\to\infty}\frac{\dfrac{\sqrt{16x^4+8}+3x^3}{3x^3}\times(3x^3)}{\dfrac{2x^2+6x+1}{2x^2}\times(2x^2)}$$

$$=\lim_{x\to\infty}\frac{\sqrt{\dfrac{16x^4+8}{9x^6}}+\dfrac{3x^3}{3x^3}}{\dfrac{2x^2+6x+1}{2x^2}}\times\frac{3x^3}{2x^2}=\lim_{x\to\infty}\frac{\sqrt{\dfrac{16}{9x^2}+\dfrac{8}{9x^6}}+1}{1+\dfrac{6}{2x}+\dfrac{1}{2x^2}}\times\frac{3x}{2}$$

$$=\frac{\sqrt{0+0}+1}{1+0+0}\times\lim_{x\to\infty}\frac{3x}{2}=\infty.$$

你一定要切实理解了后两个求解过程的每一步. 在第一个例子中, 首项来自平方根符号下的 $16x^4$; 即使当你取平方根的时候, 结果项 $4x^2$ 仍然支配了分子中的剩余部分 $(3x)$. 在第二个例子中, 占主导的则是分子中的剩余部分 $(3x^3)$. 但等一下, 你说 —— 要是它们相等会怎样呢? 例如,

$$\lim_{x\to\infty}\frac{\sqrt{4x^6-5x^5}-2x^3}{\sqrt[3]{27x^6+8x}}$$

是什么呢? 事实上, 分母并不太令人讨厌, 但还是先来看看分子吧. 在平方根符号

下, 我们有 $4x^6-5x^5$, 当 x 很大时, 它表现得就像是首项 $4x^6$. 因此, 我们应该会想 $\sqrt{4x^6-5x^5}$ 也会表现得就像是 $\sqrt{4x^6}$, 即 $2x^3$(因为 x 为正). 但问题是, 消去分子中的 $2x^3$, 似乎就没剩下什么了! 真糟糕, 应该怎么办呢?

我们使用 4.2 节中描述的技巧来求解：分子分母同时乘以分子的共轭表达式. 所以在看到首项之前, 需要做一些准备工作：

$$\lim_{x\to\infty}\frac{\sqrt{4x^6-5x^5}-2x^3}{\sqrt[3]{27x^6+8x}}=\lim_{x\to\infty}\frac{\sqrt{4x^6-5x^5}-2x^3}{\sqrt[3]{27x^6+8x}}\times\frac{\sqrt{4x^6-5x^5}+2x^3}{\sqrt{4x^6-5x^5}+2x^3}.$$

通过公式 $(a-b)(a+b)=a^2-b^2$, 可以将上式化简为

$$\lim_{x\to\infty}\frac{(4x^6-5x^5)-(2x^3)^2}{\sqrt[3]{27x^6+8x}(\sqrt{4x^6-5x^5}+2x^3)}.$$

事实上, 可以进一步整理分子, 把式子化简为

$$\lim_{x\to\infty}\frac{-5x^5}{\sqrt[3]{27x^6+8x}(\sqrt{4x^6-5x^5}+2x^3)}.$$

这就没那么糟糕了! 对于分子, 不需要再做什么, 现在来关注分母. 对于 $\sqrt[3]{27x^6+8x}$, 事实上, 可以乘以并除以首项 $27x^6$ 的立方根, 得到

$$\frac{\sqrt[3]{27x^6+8x}}{\sqrt[3]{27x^6}}\times\sqrt[3]{27x^6},$$

即

$$\frac{\sqrt[3]{27x^6+8x}}{\sqrt[3]{27x^6}}\times(3x^2).$$

当然, 可在立方根符号内合并这些项, 消去公因式, 得到

$$\sqrt[3]{\frac{27x^6+8x}{27x^6}}\times(3x^2)=\sqrt[3]{1+\frac{8}{27x^5}}\times(3x^2).$$

注意到当 $x\to\infty$ 时, 包含立方根的那部分正好趋于 1.

至于另外一项, $\sqrt{4x^6-5x^5}+2x^3$, 这里需要小心一些. 在平方根符号内有 $4x^6-5x^5$, 故其首项是 $4x^6$. 它的平方根是 $2x^3$. 现在必须把另一个 $2x^3$ 加上去, 得到分子总的“首项”, $2x^3+2x^3$, 即 $4x^3$. 来看一下这是怎么进行的吧. 将分子代以

$$\frac{\sqrt{4x^6-5x^5}+2x^3}{4x^3}\times(4x^3),$$

然后对分式进行拆分, 并把 $4x^3$ 拖进平方根符号, 它会变为 $16x^6$; 得到

$$\left(\sqrt{\frac{4x^6-5x^5}{16x^6}}+\frac{2x^3}{4x^3}\right)\times(4x^3)=\left(\sqrt{\frac{1}{4}-\frac{5}{16x}}+\frac{1}{2}\right)\times(4x^3).$$

现在, 当 $x\to\infty$ 时, 乘积的第一项就会趋于

$$\sqrt{\frac{1}{4}-0}+\frac{1}{2}=\frac{1}{2}+\frac{1}{2}=1,$$

这正是我们想要的! (注意到 $\frac{1}{4}$ 的平方根是 $\frac{1}{2}$.)

现在, 试着将所有的放在一起来求解这个问题. 由分子分母同时乘以分子的共轭表达式开始, 式子简化为

$$\lim_{x\to\infty} \frac{-5x^5}{\sqrt[3]{27x^6+8x}(\sqrt{4x^6-5x^5}+2x^3)}.$$

现在, 要在分母上使用乘除方法, 并得出

$$\lim_{x\to\infty} \frac{-5x^5}{\left(\dfrac{\sqrt[3]{27x^6+8x}}{\sqrt[3]{27x^6}}\times(3x^2)\right)\left(\dfrac{\sqrt{4x^6-5x^5}+2x^3}{4x^3}\times(4x^3)\right)}.$$

把 $-5x^5$, $3x^2$ 和 $4x^3$ 提出来, 得到

$$\lim_{x\to\infty} \frac{1}{\left(\dfrac{\sqrt[3]{27x^6+8x}}{\sqrt[3]{27x^6}}\right)\left(\dfrac{\sqrt{4x^6-5x^5}+2x^3}{4x^3}\right)}\times\frac{-5x^5}{(3x^2)(4x^3)}.$$

现在, 你所要做的只是从分子分母中消去 x^5, 并使用上面提到的论证, 来证明最后的答案是 $-5/12$. 剩下需要你做的已经不多了, 但你应该试着把以上所有的片断组合成一个完整的解.



4.5 $x \to -\infty$ 时的有理函数的极限

现在花点时间来看看形如

$$\lim_{x\to-\infty} \frac{p(x)}{q(x)}$$

的极限, 其中 p 和 q 是多项式或多项式型函数. 所有我们在一直使用的原理在这里也适用. 当 x 是一个非常大的负数时, 在任意和中, 最高次数项仍然会占主导. 此外, 当 $x \to -\infty$ 时, 只要 C 是常数, 且 n 是一个正整数, C/x^n 仍然趋于 0. (你能说出为什么吗？) 所有这些都意味着, 问题的解与之前的几乎差不多. 例如, 考虑 4.3.1 节中已经看过的那两个例子的改写



$$\lim_{x\to-\infty} \frac{x-8x^4}{7x^4+5x^3+2000x^2-6} \quad 和 \quad \lim_{x\to-\infty} \frac{(x^4+3x-99)(2-x^5)}{(18x^7+9x^6-3x^2-1)(x+1)}.$$

我所做的只是将 ∞ 改为 $-\infty$, 表明我们现在感兴趣的是, 当 x 是一个非常大的负数时, 这两个有理函数会变成什么样子. 第一个问题的解和当 $x \to \infty$ 时的解是一样的, 你只需让每个多项式分别乘以并除以其首项：

$$\lim_{x\to-\infty}\frac{x-8x^4}{7x^4+5x^3+2000x^2-6}=\lim_{x\to-\infty}\frac{\dfrac{x-8x^4}{-8x^4}\times(-8x^4)}{\dfrac{7x^4+5x^3+2000x^2-6}{7x^4}\times(7x^4)}$$

$$=\lim_{x\to-\infty}\frac{-\dfrac{1}{8x^3}+1}{1+\dfrac{5}{7x}+\dfrac{2000}{7x^2}-\dfrac{6}{7x^4}}\times\frac{-8}{7}=-\frac{8}{7}.$$

这里的要点是, 对于某个正的 n, 当 $x\to-\infty$ 时, 任何形如 C/x^n 的项都会趋于 0, 与当 $x\to\infty$ 时的情形是一样的. 另一方面, 第二个例子则不太一样; 最后一步不同于该问题之前的版本:

$$\lim_{x\to-\infty}\frac{(x^4+3x-99)(2-x^5)}{(18x^7+9x^6-3x^2-1)(x+1)}$$

$$=\lim_{x\to-\infty}\frac{\left(\dfrac{x^4+3x-99}{x^4}\times(x^4)\right)\left(\dfrac{2-x^5}{-x^5}\times(-x^5)\right)}{\left(\dfrac{18x^7+9x^6-3x^2-1}{18x^7}\times(18x^7)\right)\left(\dfrac{x+1}{x}\times(x)\right)}$$

$$=\lim_{x\to-\infty}\frac{\left(1+\dfrac{3}{x^3}-\dfrac{99}{x^4}\right)\left(-\dfrac{2}{x^5}+1\right)}{\left(1+\dfrac{9}{18x}-\dfrac{3}{18x^5}-\dfrac{1}{18x^7}\right)\left(1+\dfrac{1}{x}\right)}\times\frac{(x^4)(-x^5)}{(18x^7)(x)}$$

$$=\frac{(1+0-0)(-0+1)}{(1+0-0-0)(1+0)}\times\lim_{x\to-\infty}\frac{-x}{18}=\lim_{x\to-\infty}\frac{-x}{18}=\infty.$$

只有当在最后取极限的时候才会看到, 当 $x\to\infty$ 时和 $x\to-\infty$ 时是不同的. 现在, $-x/18$ 趋于 ∞ 而不是 $-\infty$.

还有一点需要小心. 我们之前在将因子拖进平方根符号里的时候并没有特别小心. 为了说明这一点, 试着化简 $\sqrt{x^2}$. 你会得到 x 吗? 如果不幸 x 是负的, 那你就错了. 例如, 如果平方 -2, 然后再取平方根的话, 会得到 2. 因此, 事实上, 当 x 为负时, $\sqrt{x^2}=-x$. 当你面对 $x\to-\infty$ 时的多项式型函数的极限时, 类似情况也会出现. 例如,

$$\lim_{x\to-\infty}\frac{\sqrt{4x^6+8}}{2x^3+6x+1}.$$

分母表现得就像是它的首项 $2x^3$, 但分子呢? 在平方根符号里的项 $4x^6+8$, 它表现得就像是 $4x^6$, 因此, $\sqrt{4x^6+8}$ 表现得就像是 $\sqrt{4x^6}$. 这看上去好像可以化简为 $2x^3$, 但那是不正确的! 由于 $x\to-\infty$, 我们感兴趣的是, 当 x 为负时会有什么情况发生. 这就是说, $2x^3$ 是负的, 但 $\sqrt{4x^6}$ 是正的, 所以必须将 $\sqrt{4x^6}$ 化简为 $-2x^3$. 因此, 求解过程如下:

$$\lim_{x\to-\infty}\frac{\sqrt{4x^6+8}}{2x^3+6x+1}=\lim_{x\to-\infty}\frac{\dfrac{\sqrt{4x^6+8}}{\sqrt{4x^6}}\times\sqrt{4x^6}}{\dfrac{2x^3+6x+1}{2x^3}\times(2x^3)}$$
$$=\lim_{x\to-\infty}\frac{\sqrt{\dfrac{4x^6+8}{4x^6}}}{\dfrac{2x^3+6x+1}{2x^3}}\times\frac{\sqrt{4x^6}}{2x^3}=\lim_{x\to-\infty}\frac{\sqrt{1+\dfrac{8}{4x^6}}}{1+\dfrac{6}{2x^3}+\dfrac{1}{2x^3}}\times\frac{-2x^3}{2x^3}$$
$$=\frac{\sqrt{1+0}}{1+0+0}\times(-1)=-1.$$

类似地, 在处理四次方根、六次方根等时, 你也需要同样小心. 例如,

如果 x 为负, $\sqrt[4]{x^4}=-x$.

如果用任意的偶数替换每一个 4, 结果仍然是正确的. 另一方面, 如果用一个奇数替换 4 的话, 那结果就不正确了. 例如,

对于所有的 x (正的、负的或零), $\sqrt[3]{x^3}=x$.

还有一点, 即使 $x<0$,

$$\sqrt{x^4}=x^2$$

仍然成立! 为什么呢? 因为根据定义, x^2 不可能是负的, $\sqrt{x^4}$ 也不可能是负的, 因此那里不可能有一个负号! 最后, 我们总结如下：

> 如果 $x<0$, 并且想写 $\sqrt[n]{x^{\text{某次幂}}}=x^m$, 那么需要在 x^m 之前加一个负号的唯一情形是, n 是偶的而 m 是奇的.

4.6 包含绝对值的函数的极限

有时候, 你不得不面对一些包含绝对值的函数. 试考虑极限

$$\lim_{x\to0^-}\frac{|x|}{x}.$$

为了解答此问题, 设 $f(x)=|x|/x$, 并对它检视一番. 首先, 注意到 0 不可能在函数 f 的定义域中, 因为如果 0 在其定义域中, 则分母将会是 0. 另一方面, 其他的都没问题. 我们再来看一下, 当 x 为正时 $f(x)$ 会怎样. 这时 $|x|$ 这个量就是 x, 因此, 如果 x 是任意的正数, 那么 $f(x)=1$. 另一方面, 如果 x 为负, 那么 $|x|=-x$, $f(x)=-x/x=-1$. 这就是说, $f(x)=|x|/x$ 只是 "如果 $x>0$, $f(x)=1$; 如果 $x<0$, $f(x)=-1$" 的另一种花哨说法而已. $y=f(x)$ 的图像如图 4-2 所示.

因此, 对于要求的左极限, 需要从左侧接近 $x=0$, 很明显有

$$\lim_{x\to 0^-}\frac{|x|}{x}=-1.$$

同时我们也会注意到

$$\lim_{x\to 0^+}\frac{|x|}{x}=1.$$

由于左极限和右极限不相等, 因此, 双侧极限不存在:

$$\lim_{x\to 0}\frac{|x|}{x}\text{DNE}.$$

大多数涉及绝对值的例子可以用相似的方式来解答, 即根据绝对值内部的符号, 考虑两个或更多个不同的 x 的区间. 下式是对上例的一个微小改变:

$$\lim_{x\to(-2)^-}\frac{|x+2|}{x+2}.$$

看一看这个绝对值, 就会发现, 它取决于 $x+2\geqslant 0$ 还是 $x+2<0$. 这些条件可以被重新写成 $x\geqslant -2$ 或 $x<-2$. 在第一种情况下, $|x+2|=x+2$; 而在第二种情况下, $|x+2|=-(x+2)$. 最后的结果是, 当 $x>-2$ 时, $|x+2|/(x+2)$ 等于 1; 而当 $x<-2$ 时, 它则是 -1. 事实上, $y=|x+2|/(x+2)$ 的图像就是 $y=|x|/x$ 的图像向左平移两个单位得到的, 如图 4-3 所示.

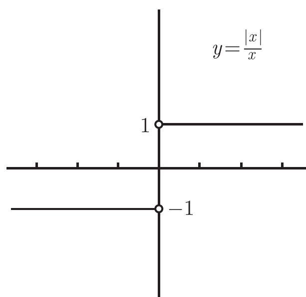


图　4-2

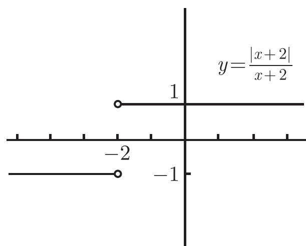


图　4-3

这就是说, 要求的左极限等于 -1 (同时, 右极限是 1, 故双侧极限不存在).

第 5 章　连续性和可导性

一般而言, 函数的图像只有一点比较特殊：它必须满足垂线检验. 这并没有要求特别多. 图像可以散落四处：这里有一部分, 那里有一条垂直渐近线, 或者随心所欲地在各处散落任意个不连续的点. 所以现在我们想要看看, 如果对函数图像要求略微多一点会发生什么：我们将要讨论两种类型的**光滑性**. 首先是连续性, 直觉告诉我们, 连续函数的图像必须能一笔画成. 其次是可导性, 直觉上, 在可导函数的图像中不会出现尖角. 在这两种情形中, 我们都将深入地讨论其定义, 并了解满足这些特殊要求的函数具有的一些性质. 详细地说, 以下是我们将在本章中所要研究的内容：

- 在一点处及在一个区间上连续;
- 连续函数的一些例子;
- 连续函数的介值定理;
- 连续函数的最大值与最小值;
- 位移、平均速度和瞬时速度;
- 切线和导数;
- 二阶导和高阶导;
- 连续性和可导性的关系.

5.1　连　续　性

我们先从一个函数是连续的, 这到底意味着什么开始. 正如我上面所说, 直觉上, 可以一笔画出连续函数的图像. 这对于像 $y = x^2$ 这样的函数来说没有问题, 因为整个图像在一块; 但对于像 $y = 1/x$ 这样的函数, 这就有一点儿不公平了. 要不是在 $x = 0$ 处有一条垂直渐近线, 把图像分成了两部分, 它的图像本来可以是在一块的. 事实上, 如果 $f(x) = 1/x$, 那么可以说, 除了在 $x = 0$ 外, f 处处连续. 因此, 必须理解在一点处连续是什么意思. 然后, 考虑在更大的区域上, 比如区间上的连续性.

5.1.1　在一点处连续

我们以一个函数 f 和在 x 轴上其定义域中的点 a 开始. 当我们画 $y = f(x)$ 的图像时, 想要在通过图像上的点 $(a, f(a))$ 时不提起笔. 如果在其他地方必须提起笔的话, 那也不要紧, 只要在 $(a, f(a))$ 的附近不提起笔就行了. 这意味着, 我们想要

一连串点 $(x, f(x))$ 变得越来越接近 (事实上是任意地接近) 于点 $(a, f(a))$. 换句话说, 当 $x \to a$ 时, 需要 $f(x) \to f(a)$. 没错, 女士们, 先生们, 我们这里面对的是极限问题. 现在可以给出一个恰当的定义：

如果 $\lim\limits_{x\to a} f(x) = f(a)$, 函数 f 在点 $x = a$ 处连续.

当然, 为了让前面的等式有意义, 等号两边必须都是有定义的. 如果极限不存在, 那么 f 在点 $x = a$ 处不连续, 而如果 $f(a)$ 不存在, 那么你彻底完蛋了：那里甚至都没有一个点 $(a, f(a))$ 可以让你通过! 因此, 可以对定义进行更精确一些的描述, 并明确地要求以下三条成立：

(1) 双侧极限 $\lim\limits_{x\to a} f(x)$ 存在 (并且是有限的);

(2) 函数在点 $x = a$ 处有定义, 即 $f(a)$ 存在 (并且是有限的);

(3) 以上两个量相等, 即

$$\lim_{x\to a} f(x) = f(a).$$

让我们来看看, 如果任意一条性质不满足, 那会怎么样. 考虑图 5-1.

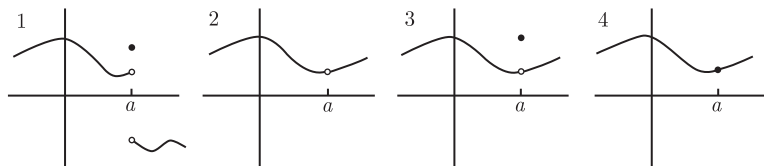


图　5-1

在标号为 1 的图中, 在 $x = a$ 处的左极限和右极限不相等, 则双侧极限不存在, 所以函数在点 $x = a$ 处不连续. 在标号为 2 的图中, 左极限和右极限都存在且是有限的, 并且左右极限相等, 故双侧极限存在; 然而, 函数在点 $x = a$ 处无定义, 因此, 函数在点 $x = a$ 处不连续. 在标号为 3 的图中, 双侧极限也存在, 函数在点 $x = a$ 处有定义, 但极限值和函数值不相等, 再一次地, 函数在点 $x = a$ 处一次不连续. 另一方面, 在标号为 4 的图中, 由于双侧极限在点 $x = a$ 处存在, $f(a)$ 存在, 并且极限值和函数值相等, 因此, 函数的确在点 $x = a$ 处连续. 顺便说一下, 前三个图中的函数在点 $x = a$ 处有一个**不连续点**.

5.1.2　在一个区间上连续

我们已经知道函数在一个单点上连续的定义了. 现在来把该定义扩展一下, 如果函数在区间 (a, b) 上的每一点都连续, 那么它在该区间上连续. 注意到 f 实际上没有必要在端点 $x = a$ 或 $x = b$ 上连续. 例如, 如果 $f(x) = 1/x$, 那么 f 在区间 $(0, \infty)$ 上连续, 即使 $f(0)$ 无定义. 该函数在区间 $(-\infty, 0)$ 上也连续, 但在区间 $(-2, 3)$ 上不连续, 因为 0 位于此区间内, 而 f 在那里不连续.

对于形如 $[a,b]$ 的区间又如何呢? 对此我们不得不稍微灵活些. 例如, 图 5-2 是函数在其定义域 $[a,b]$ 上的图像; 我们想说它在 $[a,b]$ 上连续. 但问题是, 双侧极限在端点 $x=a$ 和 $x=b$ 处不存在: 在点 $x=a$, 只有一个右极限; 而在点 $x=b$, 只有一个左极限. 不过没有关系, 只需利用端点处适当的单侧极限来略微修改一下定义. 因此, 我们说函数 f 在 $[a,b]$ 上连续, 如果

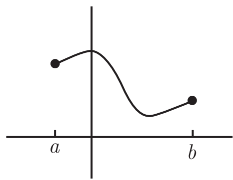


图 5-2

(1) 函数 f 在 (a,b) 中的每一点都连续;

(2) 函数 f 在点 $x=a$ 处**右连续**; 即, $\lim\limits_{x\to a^+} f(x)$ 存在 (且有限), $f(a)$ 存在, 并且这两个量相等; 以及

(3) 函数 f 在点 $x=b$ 处**左连续**; 即, $\lim\limits_{x\to b^-} f(x)$ 存在 (且有限), $f(b)$ 存在, 并且这两个量相等.

最后, 如果函数在其定义域中的所有的点都连续, 我们就说它是**连续的**. 如果函数的定义域包括一个带有左端点和/或右端点的区间, 那么在那里需要函数的单侧连续性.

5.1.3 连续函数的一些例子

很多的常见函数都是连续的. 例如, 每一个多项式都是连续的. 这看起来好像不太好证明, 因为有很多不同的多项式, 但事实上并不是那么难证明. 首先, 让我们证明定义为 $f(x)=1$ 的常数函数 f, 对于所有的 x, 在任意一点 a 处都连续. 也就是说, 需要证明

$$\lim_{x\to a} f(x) = f(a).$$

由于对于任意的 x 都有 $f(x)=1$, 并且 $f(a)=1$, 这意味着需要证明

$$\lim_{x\to a} 1 = 1.$$

显然上式成立, 因为所有的一切都不依赖于 x 和 a. 现在, 设 $g(x)=x$. g 是连续的吗? 这时需要证明

$$\lim_{x\to a} g(x) = g(a).$$

由于 $g(x)=x$ 且 $g(a)=a$, 这就将问题简化为证明

$$\lim_{x\to a} x = a.$$

显然上式也成立: 当 $x\to a$ 时, 当然会有 $x\to a$! 现在只需观察可知, 一个连续函数的常数倍是连续的; 此外, 如果对两个连续函数做加法、减法、乘法或复合, 会得到另一个连续函数 (更多详情请参见附录 A 的 A.4.1 节). 当用一个连续函数除以另一个连续函数的时候, 这几乎也一样成立: 除了分母为零的点外, 商函数处处连续. 例如, 除了在 $x=0$ 处, $1/x$ 在其他各处都是连续的, 因为我们已经看到分子分

母同为 x 的连续函数.

不管怎样, 让我们回到多项式. 因为 $g(x) = x$ 是 x 的连续函数, 可以让 g 和它自己相乘, 看到 x^2 也是 x 的连续函数. 你想要多少个 x 和它自己相乘都可以, 这样可以证明 x 的任意次幂 (作为 x 的函数) 的连续性. 然后, 可以乘以常数系数, 并将不同次幂相加在一起, 得到任意一个多项式 —— 并且每一个仍然是连续的!

结果证明, 所有的指数函数和对数函数都是连续的, 同样所有的三角函数也是如此 (除了在它们的渐近线上). 我们暂且接受这一点, 后面的 5.2.11 节将会解释其中的原因. 同时, 我想让你来看一个更奇异的函数. 考虑函数 f, 其定义为 $f(x) = x\sin(1/x)$. 在 3.6 节有过它的图像 (至少是当 $x > 0$ 时的图像). 其实把图像扩展到 $x < 0$ 是很容易的, 因为 f 是一个偶函数. 为什么呢? 记得 $\sin(x)$ 是 x 的奇函数, 于是有

$$f(-x) = (-x)\sin\left(\frac{1}{-x}\right) = (-x)\left(-\sin\left(\frac{1}{x}\right)\right) = x\sin\left(\frac{1}{x}\right) = f(x).$$

因此, f 的确是偶函数, 从而以 y 轴为镜子反射之前的图像, 就可以得到 f 的图像 (图 5-3 只显示了在定义域 $-0.3 < x < 0.3$ 上的图像).

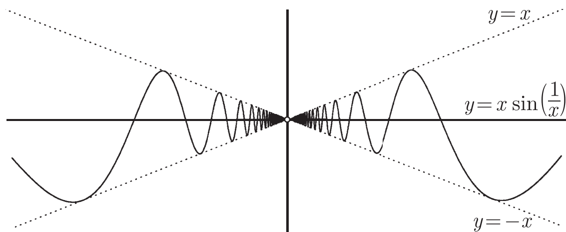


图　5-3

现在来考虑一下这个函数的连续性. 作为 x 的函数, 我们已经知道, 除了在 $x = 0$ 处, $1/x$ 在其他各处都是连续的. 现在, 我们将它与正弦函数作复合, 得到的函数依然是连续的, 并且可以看到, 除了在 $x = 0$ 处, $\sin(1/x)$ 在其他各处也都是连续的. 现在, 你只需要用 x (这显然是 x 的连续函数!) 和 $\sin(1/x)$ 相乘就可以看到除了在 $x = 0$ 处外, f 在其他各处都是连续的.

那么在 $x = 0$ 处发生了什么呢? 显然, f 在 $x = 0$ 不连续, 因为它在那里甚至都没有定义 (图像上这里有一个洞). 让我们来定义一个函数 g 如下, 将这个洞堵上:

$$g(x) = \begin{cases} x\sin\left(\dfrac{1}{x}\right) & \text{如果 } x \neq 0, \\ 0 & \text{如果 } x = 0. \end{cases}$$

因此, 除了在 $x = 0$ (此时 g 等于 0, 而 f 无定义) 外, $g(x) = f(x)$. 因此, g 必然是处处连续的, 而 f 除 $x = 0$ 外处处连续. 现在需要来看看在 $x = 0$ 处发生了什么.

由于 $g(0)$ 有定义, 这就有了希望. 此外, 可以使用 3.6 节的三明治定理来证明

$$\lim_{x\to 0^+} g(x) = \lim_{x\to 0^+} x\sin\left(\frac{1}{x}\right) = 0.$$

通过对称性 (或再次使用三明治定理), 可以看到左极限也等于 0. 事实上, 双侧极限也为 0:

$$\lim_{x\to 0} g(x) = \lim_{x\to 0} x\sin\left(\frac{1}{x}\right) = 0.$$

因此, 就证明了

$$\lim_{x\to 0} g(x) = g(0),$$

因为等号两边都存在且等于 0. 这意味着, g 在 $x=0$ 处实际上是连续的, 尽管它是一个分段函数的形式.

我们差不多已经准备好, 可以来看看两个涉及连续性的很好的事实. 不过首先, 我想回到第 4 章开始时曾讲到的一点. 当时所举的第一个例子是

$$\lim_{x\to -1} \frac{x^2-3x+2}{x-2},$$

将 $x=-1$ 代入上式求解得到结果为 -2. 为什么可以这样做? 这似乎与之前所说, 上述极限的值与在 $x=-1$ 处发生的情况无关, 仅仅与在 $x=-1$ 附近的情况有关这一点相矛盾. 这里就轮到连续性派上用场了: 它将 “附近的” 与 “在” 联系了起来. 特别是, 如果令 $f(x) = \left(x^2-3x+2\right)/(x-2)$, 那么由于分子和分母都是多项式, 除了在分母为 0 的点外, f 是处处连续的. 也就是说, 除了在 $x=2$ 处, f 是处处连续的. 因此, f 在 $x=-1$ 上是连续的, 这就意味着,

$$\lim_{x\to -1} f(x) = f(-1).$$

用其定义替换 f, 有

$$\lim_{x\to -1} \frac{x^2-3x+2}{x-2} = \frac{(-1)^2-3(-1)+2}{(-1)-2} = -2.$$

这就是完整的解. 在实践中, 很少有数学家会不厌其烦地把这些细节都写出来, 但这样做会有助于你理解你在做什么!

5.1.4 介值定理

知道一个函数是连续的会有很多好处. 我们将看看其中两个好处. 第一个被称为**介值定理**. 其基本思想是: 假设一个函数 f 在一个闭区间 $[a,b]$ 上连续. 此外, 假设 $f(a)<0$ 且 $f(b)>0$. 因此, 在 $y=f(x)$ 的图像上, 点 $(a,f(a))$ 位于 x 轴的下方, 而点 $(b,f(b))$ 位于 x 轴的上方, 如图 5-4 所示.

现在, 如果必须用一条曲线 (当然它要满足垂线检验) 来连接这两个点, 并且不

允许抬起笔来, 直觉上显然有, 你的笔将与 x 轴上 a 和 b 之间的某处至少相交一次. 交点也许在 a 的附近或 b 的附近, 或者在 a 和 b 中间的某处, 但必须相交至少一次. 这就是说, x 轴截距在 a 和 b 之间的某处. 在这里, 函数 f 在区间 $[a,b]$ 上的每一点都是连续的, 这一点至关重要; 我们来看看哪怕 f 仅仅在一点处不连续会怎样, 如图 5-5 所示.

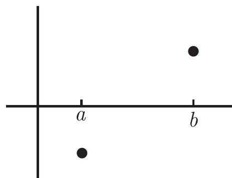


图　5-4

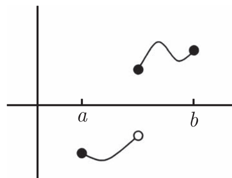


图　5-5

不连续点让函数在 x 轴上发生跳跃而不通过 x 轴. 因此, 需要在整个区域 $[a,b]$ 上的连续性. 这也适用于从 x 轴上方开始并在 x 轴下方结束的情况, 即如果 $f(a)>0$ 且 $f(b)<0$, 并且 f 在 $[a,b]$ 上的每一点都连续, 那么在 $[a,b]$ 上的某处, 必定会有一个 x 轴截距. 由于 x 轴截距意味着 $f(c)=0$, 可以表述介值定理如下：

> **介值定理：** 如果 f 在 $[a,b]$ 上连续, 并且 $f(a)<0$ 且 $f(b)>0$, 那么在区间 (a,b) 上至少有一点 c, 使得 $f(c)=0$. 代之以 $f(a)>0$ 且 $f(b)<0$, 同样成立.

该定理的证明请参见附录 A 中的 A.4.2 节. 现在来看一些如何应用此定理的例子. 首先, 假设要证明多项式 $p(x)=-x^5+x^4+3x+1$ 在 $x=1$ 和 $x=2$ 之间有一个 x 轴截距. 你只需注意到, 由于它是一个多项式, 所以 p 是处处连续的 (包含 $[1,2]$); 此外, 计算 $p(1)=4>0$ 且 $p(2)=-9<0$. 由于 $p(1)$ 和 $p(2)$ 的符号相反, 且 p 在 $[1,2]$ 上连续, 我们知道在区间 $(1,2)$ 上至少存在一点 c 使得 $p(c)=0$. 数 c 就是多项式 p 的一个 x 轴截距.

接着是一个稍微难一点的例子. 如何证明方程 $x=\cos(x)$ 有一个解呢？不需要求出解来, 只需要证明存在一个解. 可以先在同一坐标轴上画出 $y=x$ 和 $y=\cos(x)$ 的图像. 如果这样做了, 就会发现图像的交点的 x 轴坐标在 $\pi/4$ 附近. 不过这样的图像式论证, 虽然不无说服力, 但对于一个数学证明来说, 还远远不够. 那么如何能够做得更好呢？

第一步是使用一个小窍门：**将所有表达式放到等号左边**. 因此, 我们试着来求解 $x-\cos(x)=0$, 而不是求解 $x=\cos(x)$. 现在, 设 $f(x)=x-\cos(x)$. 如果可以证明存在数 c 使得 $f(c)=0$ 的话, 任务就算完成了. 来检验一下这是否说得通：如果 $f(c)=0$, 那么 $c-\cos(c)=0$, 因此 $c=\cos(c)$, 于是就找到了方程 $x=\cos(x)$

的一个解, 它就是 $x = c$.

现在, 该使用介值定理了. 我们需要找到两个数 a 和 b, 使得 $f(a)$ 和 $f(b)$ 其中一个是负的而另一个是正的. 由于从图像中可知答案会在 $\pi/4$ 附近, 我们将保守地选取 $a = 0$ 和 $b = \pi/2$. 来检验一下 $f(0)$ 和 $f(\pi/2)$ 的值吧. 首先, $f(0) = 0 - \cos(0) = 0 - 1 = -1$, 它是负的; 其次, $f(\pi/2) = \pi/2 - \cos(\pi/2) = \pi/2 - 0 = \pi/2$, 它是正的. 由于 f 是连续的 (它是两个连续函数的差), 根据介值定理可以得出, 在区间 $(0, \pi/2)$ 上存在某个数 c 使得 $f(c) = 0$, 于是证明了 $x = \cos(x)$ 有一个解. 我们不知道解在哪里, 也不知道会有多少解, 只是知道在区间 $(0, \pi/2)$ 上至少有一个解. (注意, 解实际上不是 $\pi/4$! 事实上, 不可能找到一个有关解的很好的表达.)

这里有一个稍有不同的变体. 到目前为止, 都是规定 $f(a) < 0$ 且 $f(b) > 0$ (或反过来), 然后得出结论, 在 (a, b) 上存在一点 c 使得 $f(c) = 0$. 然而现在, 可以用任意数 M 来替换 0, 且结果依然成立. 因此, 假设 f 在 $[a, b]$ 上连续; 如果 $f(a) < M$ 且 $f(b) > M$ (或反过来), 那么在 (a, b) 上存在一点 c 使得 $f(c) = M$. 例如, 如果 $f(x) = 3^x + x^2$, 那么方程 $f(x) = 5$ 有解吗? 显然 f 是连续的; 我们也可以猜出解在 0 和 2 之间, 这样会有 $f(0) = 1$ 和 $f(2) = 13$. 由于数 1 和 13 夹着目标数 5(一个小一点而另一个大一点), 介值定理告诉我们, 对于 $(0, 2)$ 上的某个 c 有 $f(c) = 5$.

这就是说, $f(x) = 5$ 确实有解. 现在试着以一个新的函数 g 来重新做一遍, 其定义为 $g(x) = 3^x + x^2 - 5$. 可以看出, 如果 $f(x) = 5$ 有一个解是 c, 那么 c 也是 $g(x) = 0$ 的解. 由于 $g(0) < 0$ 且 $g(2) > 0$, 你可以使用先前的方法而不是上面的变体! 事实上, 变体并没有给我们提供任何新的东西, 它只是有时候会让生活变得更简单些.

5.1.5 一个更难的介值定理例子

最后一个例子：证明任意的奇数次多项式至少有一个根. 这就是说, 令 p 是一个奇数次多项式, 我断言, 至少有一个数 c 使得 $p(c) = 0$. (这对于偶数次多项式不成立. 例如, 二次的 $x^2 + 1$ 没有根, 其图像和 x 轴不相交.) 可是如何来证明我的断言呢?

事实上, 这里的关键可以追溯到 4.3 节. 在那里, 如果 $p(x)$ 是任意的多项式, 且其首项为 $a_n x^n$, 那么

$$\lim_{x\to\infty} \frac{p(x)}{a_n x^n} = 1 \quad \text{且} \quad \lim_{x\to-\infty} \frac{p(x)}{a_n x^n} = 1.$$

因此, 当 x 变得非常大时, $p(x)$ 和 $a_n x^n$ 会相对地非常接近 (它们的比值接近于 1). 这意味着, 它们至少有相同的符号! 不可能是一个负一个正, 否则它们的比值为负,

而不是接近于 1. 当 x 是一个非常大的负数时, 情况也是如此.

因此, 假设 A 是一个很大的负数, 使得 $p(A)$ 和 a_nA^n 有相同的符号. 此外, 选取一个非常大的正数 B, 使得 $p(B)$ 和 a_nB^n 有相同的符号. 现在, 比较一下 a_nA^n 和 a_nB^n 的符号. 由于 n 是一个奇数, 它们的符号一定相反! 一个为正而另一个为负. 例如, 如果 $a_n > 0$, 那么 a_nB^n 为正且 a_nA^n 为负. (这只有当 n 是奇数时才成立：如果 n 是偶数, 那么这两个量均为正.) 因此, 有

$$p(A) \overset{\text{符号相同}}{\longleftrightarrow} a_nA^n \overset{\text{符号相反}}{\longleftrightarrow} a_nB^n \overset{\text{符号相同}}{\longleftrightarrow} p(B).$$

所以 $p(A)$ 和 $p(B)$ 的符号相反. 由于 p 是一个多项式, 它是连续的, 于是根据介值定理, 在 A 和 B 之间有一个数 c, 使得 $p(c) = 0$. 这就是说, p 有一个根, 尽管不知道它在哪儿. 这也没办法, 毕竟不知道 p 是什么样子的多项式, 只知道它是奇数次的.

5.1.6　连续函数的最大值和最小值

接着来看知道一个函数是连续的所带来的第二个好处. 假设有一个函数 f, 它在闭区间 $[a,b]$ 上连续. (这里区间的两个端点都是闭的非常重要.) 这意味着, 可以拿笔放在点 $(a,f(a))$ 上, 由此出发, 笔不离纸地画一条曲线, 并结束于点 $(b,f(b))$. 这里的问题是, 能画多高? 换句话说, 这条曲线能够达到的高度有限度吗? 回答是肯定的, 一定有一个最高点, 尽管曲线可以多次达到最高点.

用符号表达即为, 定义在区间 $[a,b]$ 上的函数 f 在 $x = c$ 处有一个**最大值**, 如果 $f(c)$ 是 f 在整个区间 $[a,b]$ 上的最大值. 即对于区间上所有的 x, $f(c) \geqslant f(x)$. 这里我试图传递的基本思想是, $[a,b]$ 上的连续函数在区间 $[a,b]$ 上有最大值. 对于类似的问题, “能画多低”, 我们也有同样的说法, 即 f 在 $x = c$ 处有一个**最小值**, 如果 $f(c)$ 是 f 在整个区间 $[a,b]$ 上的最小值. 即对于 $[a,b]$ 上所有的 x, $f(c) \leqslant f(x)$. 再一次地, 区间 $[a,b]$ 上的任何连续函数在该区间上都有最小值. 这些事实构成一个定理, 有时被称作**最大值与最小值定理**, 它可以陈述如下：

> **最大值与最小值定理**：如果 f 在 $[a,b]$ 上连续, 那么 f 在 $[a,b]$ 上至少有一个最大值和一个最小值.

图 5-6 是一些关于 $[a,b]$ 上的连续函数及其最大值与最小值的例子.

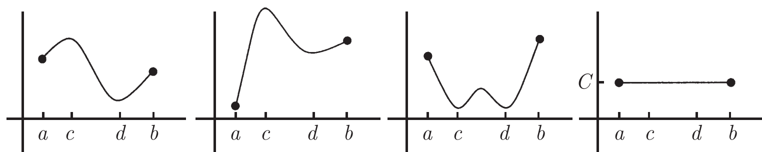


图　5-6

在第一幅图中, 函数在 $x = c$ 处取得最大值并在 $x = d$ 处取得最小值. 在第二

幅图中, 函数在 $x = c$ 处取得最大值而在左端点 $x = a$ 处取得最小值. 在第三幅图中, 最大值在 $x = b$ 处, 而最小值在 $x = c$ 和 $x = d$ 上. 这是可以接受的 (允许有多个最小值, 只要至少有一个). 最后, 第四幅图展示了一个常数函数, 它是连续的; 事实上, 由于该函数绝不会高于或低于常数 C, 所以区间 $[a, b]$ 中的每一个点既是最大值也是最小值.

那么为什么需要函数 f 是连续的? 并且, 为什么不能是一个像 (a, b) 那样的开区间? 图 5-7 显示了一些潜在的问题.

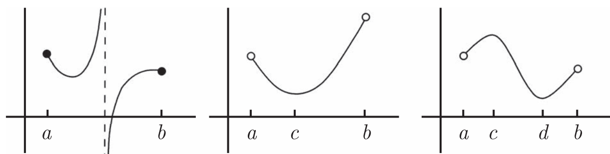


图 5-7

在第一幅图中, 函数 f 在区间 $[a, b]$ 的中间有一条渐近线, 它当然会产生一个不连续点. 该函数没有最大值, 它只会在渐近线的左侧无限上升. 类似地, 它也没有最小值, 因为它会在渐近线的右侧无限下降.

第二幅图涉及一个更微妙的情况. 这里函数只在开区间 (a, b) 上连续. 显然该函数在 $x = c$ 处有一个最小值, 但它的最大值是什么呢? 你或许会想它出现在 $x = b$ 处, 但再想想看. 该函数在 $x = b$ 处没有定义! 因此, 它不可能在那里有一个最大值. 如果该函数有一个最大值, 那么它一定在 b 附近的某处. 事实上, 你想要的是一个小于 b 并接近于 b 的数. 很不幸, 没有这样的数! 无论你想到一个多么接近于 b 的数, 你总是可以取该数与 b 的平均数得到另一个更接近于 b 的数. 因此, 该函数没有最大值. 这说明, 为了确保可以使用最大值与最小值定理, 连续性区间必须是闭的.

当然, 即使区间不是闭的, 该定理的结论也可能会成立. 例如, 在上面的第三幅图中, 函数只在开区间 (a, b) 上连续, 但它仍然在 $x = c$ 处有一个最大值并在 $x = d$ 处有一个最小值. 但这只是一个幸运情况. 如果你知道函数在区间 $[a, b]$ 上连续, 你只能仰赖定理来确保最大值与最小值的存在性.

5.2 可 导 性

我们已经花了一些时间来学习连续性. 现在该来看看函数能够具有的另一种光滑性 —— 可导性. 这实质上意味着函数有导数. 因此, 我们会花相当一部分时间来研究导数. 发展微积分的最初灵感之一来自试图去理解运动物体的速度、距离和时间的关系. 因此, 让我们从那里开始, 之后再回到函数.

5.2.1 平均速率

想象一下, 在高速路上给一辆汽车拍照. 曝光时间非常短, 因此图像并不模糊 —— 你甚至不能分辨那辆车是不是在动. 现在, 我问你: 拍照时汽车的运动速度有多快? 你说, 没问题, 只需使用经典公式

$$\text{速率} = \frac{\text{距离}}{\text{时间}}.$$

但问题是, 照片无法告诉你距离 (那辆车没有动) 或时间 (照片实质上是捕捉了一瞬间). 因此, 你无法回答我的问题.

嗯, 但如果我告诉你, 拍照之后的一分钟, 汽车行驶了一英里呢? 这时你就可以使用以上公式来计算了, 汽车一分钟开了一英里, 速率是 60 英里/小时. 但仍旧, 你如何知道汽车在那一分钟里的速率是一样的呢? 在那一分钟里, 它可能会有多次的加速和减速. 你不知道在那一分钟的开始时刻它究竟开得有多快. 事实上, 上述公式并不精确: 等号左边应该称为**平均速率**, 因为那是我们所能知道的全部.

好吧, 看你可怜, 我再告诉你, 在第一个 10 秒钟, 汽车行驶了 0.25 英里. 现在, 你可以使用该公式来计算, 在第一个 10 秒钟内的平均速率是 1.5 英里/分钟或 90 英里/小时. 这有点帮助, 但在这 10 秒钟里汽车仍旧可能改变过速率, 因此我们仍然不知道在这段时间的开始时刻它开得有多快. 不过速率也不可能跟 90 英里/小时差太多, 毕竟在这么短的时间里, 汽车只可能加速或减速这么多.

如果知道在拍照后的一秒钟里汽车走了多远, 那将会更好, 但这仍旧还不够. 甚至 0.0001 秒都可能足以让汽车改变速率, 尽管变化不会太大. 如果你感到我们是在取极限的话, 那你想得没错. 不过, 我们首先需要看一看速度的概念.

5.2.2 位移和速度

想象一下, 汽车在一条长直的高速路上行驶. 公路上的里程标志牌有点奇怪: 某个点上是 0 标志, 在其左侧, 标志始于 −1 并且变得越来越负; 在其右侧, 一切一如平常. 事实上, 整个情形看上去就像图 5-8.

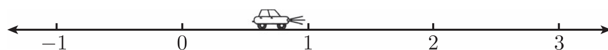


图 5-8

假设汽车始于 2 英里处并直接驶向 5 英里处, 那么它行驶的距离是 3 英里. 但如果它是始于 2 英里处但向左行驶到了 −1 英里处, 它行驶的距离也是 3 英里. 我们想要区分这两种情形, 因此我们将使用**位移**来代替距离. 位移公式就是:

$$\text{位移} = \text{终点位置} - \text{初始位置}.$$

如果汽车从位置 2 驶到位置 5, 那么位移是 $5 - 2 = 3$ 英里. 但如果是从位置 2 驶到了位置 −1, 那么位移是 $(-1) - 2 = -3$ 英里. 因此, 和距离不一样, 位移可以是

负的. 事实上, 如果位移是负的, 那么汽车将终止于它初始位置的左侧.

距离和位移的另外一个重要区别就是, 位移仅仅涉及终点和初始位置, 汽车在行驶过程中的情况是无关紧要的. 如果它从 2 走到 11, 然后又返回到 5, 距离是 $9+6=15$ 英里, 但总位移仍然只是 3 英里. 而如果它从 2 走到 -4 然后又返回到 2, 位移实际上是 0 英里, 尽管距离是 12 英里. 然而, 如果汽车只向一个方向行驶, 没有后退的话, 那么距离就是位移的绝对值.

正如我们在上一节看到的, 平均速率是行使距离除以行驶时间. 如果你用位移来代替距离, 你会得到**平均速度**. 也就是,

$$\text{平均速度} = \frac{\text{位移}}{\text{时间}}.$$

同样, 速度可以是负的, 而速率必定是非负的. 如果在一定的时间段内, 汽车有一个负的平均速度, 那么它终止于初始位置的左侧. 而如果在一定的时间段内平均速度是 0, 那么汽车终止于它的初始位置. 注意到, 在这种情况下, 汽车或许有一个很高的平均速率, 尽管其平均速度为 0! 一般而言, 就像位移, 如果汽车沿着一个方向行驶, 那么平均速率就是平均速度的绝对值.

5.2.3 瞬时速度

现在, 我们用速度来重新考察一下前面提到的重要问题：在给定的瞬间, 如何测量汽车的速度？如前所述, 基本思想就是, 在始于拍照时刻并变得越来越小的时间段上, 求汽车的平均速度. 下面就是如何用符号来表达这个思路.

令 t 是我们关心的时刻. 例如, 如果全程始于下午两点, 你可能决定要以秒表记, 并用 0 表示开始时间. 那种情况下, 如果拍照时间是下午两点零三分, 那么你将取 $t=180$. 不管怎样, 假设 u 是 t 之后很近的时刻. 我们写 $v_{t\leftrightarrow u}$ 表示汽车在始于时间 t 终止于时间 u 的时间段上的平均速度. 现在, 让 u 越来越靠近 t. 多近呢？能有多近就多近! 而这正是轮到极限登场的地方. 事实上,

$$\text{在时刻 } t \text{ 的瞬时速度} = \lim_{u\to t^+} v_{t\leftrightarrow u}.$$

不过, 为什么要忽略在时刻 t 之前的细节呢？通过允许 u 在 t 之前, 我们可以让以上定义变得更一般一些. 然后, 我们可以用双侧极限替换右极限：

$$\text{在时刻 } t \text{ 的瞬时速度} = \lim_{u\to t} v_{t\leftrightarrow u}.$$

现在需要更多的公式. 假设知道在高速路上汽车在任意时刻的准确位置. 特别是, 假设在时刻 t, 汽车的位置是 $f(t)$. 这就是说, 令

$$f(t) = \text{汽车在时刻 } t \text{ 的位置}.$$

现在就可以准确地计算平均速度 $v_{t\leftrightarrow u}$ 了：

$$v_{t\leftrightarrow u} = \frac{\text{在时刻 } u \text{ 的位置} - \text{在时刻 } t \text{ 的位置}}{u-t} = \frac{f(u)-f(t)}{u-t}.$$

注意到分母 $u-t$ 是所涉及时间段的长度 (如果 u 在 t 之后的话[①]). 不管怎么说, 现在来取 $u \to t$ 时的极限:

$$\text{在时刻 } t \text{ 的瞬时速度} = \lim_{u\to t} \frac{f(u)-f(t)}{u-t}.$$

当然, 在以上极限中, 不能只是用 $u=t$ 作替换, 因为那样的话, 会得到 0/0 的不定式. 你现在还是要使用极限形式.

再来看一个稍有变化的变体. 我们定义 $h=u-t$. 由于 u 非常靠近 t, 两时刻的差值 h 一定非常小. 确实, 当 $u \to t$ 时, 可以看到 $h \to 0$. 如果在上述极限中作如此替换的话, 由于 $u=t+h$, 也会有

$$\text{在时刻 } t \text{ 的瞬时速度} = \lim_{h\to 0} \frac{f(t+h)-f(t)}{h}.$$

该公式和前一个公式没有实质性差别, 只是写法不同而已.

让我们来看一个小的例子. 假设处于静止状态的汽车从 7 英里标志处向右开始加速, 并设此时刻 $t=0$ 小时. 结果表明, 汽车在时刻 t 的位置好像是 $15t^2+7$(这里的数 15 取决于加速度). 暂且不去担心为什么会如此, 让我们设 $f(t)=15t^2+7$, 并看看是否可以求出汽车在任意时刻 t 的速度.

使用上述公式有

$$\begin{aligned}\text{在时刻 } t \text{ 的瞬时速度} &= \lim_{h\to 0} \frac{f(t+h)-f(t)}{h} \\ &= \lim_{h\to 0} \frac{(15(t+h)^2+7)-(15t^2+7)}{h}.\end{aligned}$$

现在展开 $(t+h)^2=t^2+2th+h^2$, 并进一步化简, 看到上述表达式变为

$$\lim_{h\to 0} \frac{15t^2+30th+15h^2+7-15t^2-7}{h} = \lim_{h\to 0} \frac{30th+15h^2}{h} = \lim_{h\to 0}(30t+15h).$$

在最后一步, 从分母中消去了 h, 这非常好, 因为是它造成了所有的麻烦. 现在, 就可以将 $h=0$ 代入并看到

$$\text{在时刻 } t \text{ 的瞬时速度} = \lim_{h\to 0}(30t+15h) = 30t.$$

因此, 在时刻 0, 汽车的速度是 $30\times 0=0$ 英里/小时 —— 汽车处于静止状态. 半小时之后, 在时刻 $t=1/2$, 它的速度是 $30\times 1/2=15$ 英里/小时. 一小时之后, 速度是 30 英里/小时. 事实上, 在时刻 t 的速度是 $30t$, 这个事实告诉我们, 汽车行驶得越来越快, 每小时速度增加 30 英里/小时. 也就是说, 汽车以 30 英里每二次方小时加速.

5.2.4 速度的图像阐释

是时候来看看图像了. 再次假设 $f(t)$ 代表汽车在时刻 t 的位置. 如果想要在

① 如果 u 在 t 之前, 那么分母应该是 $t-u$, 分子应该是 $f(t)-f(u)$, 因此无论怎样都没问题!

特定时刻 t 的瞬时速度, 需要选取一个靠近 t 的时刻 u. 让我们来画一下 $y = f(t)$ 的图像, 并标注位置 $(t, f(t))$ 和 $(u, f(u))$ 以及过这两点的直线, 如图 5-9 所示.

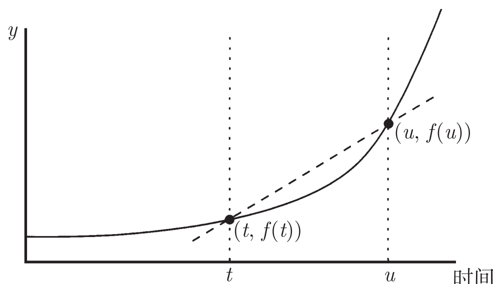


图 5-9

该直线的斜率由公式

$$斜率 = \frac{f(u) - f(t)}{u - t}$$

给出, 这正好就是上一节中平均速度 $v_{t\leftrightarrow u}$ 的公式. 因此就有了在 t 到 u 时间段上平均速度的图像阐释：在位置与时间的图像上, 它就是连接点 $(t, f(t))$ 和 $(u, f(u))$ 的直线的斜率.

让我们来试着给瞬时速度找一个类似的阐释. 我们需要取 u 趋于 t 时的极限, 因此要重复几次上述图像, 每一次 u 会越来越接近固定值 t, 如图 5-10 所示.

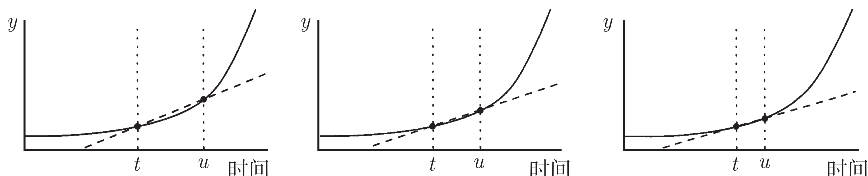


图 5-10

这些直线看上去好像越来越接近点 $(t, f(t))$ 处的切线. 由于瞬时速度是这些直线在 $u \to t$ 时的极限, 于是, 瞬时速度就等于通过点 $(t, f(t))$ 的切线的斜率. 看起来需要对切线有更好的了解 ……

5.2.5 切线

假设在某个函数 f 的定义域上选取一点 x, 那么点 $(x, f(x))$ 位于 $y = f(x)$ 的图像上. 我们想要试着画一条通过该点并与该曲线相切的直线, 即要找到一条切线. 直观上, 这意味着要找的直线刚好掠过该曲线的点 $(x, f(x))$. 切线不是只能与曲线仅相交一次! 例如, 图 5-11 中通过点 $(x, f(x))$ 的切线与曲线还有第二次相交, 这不成问题.

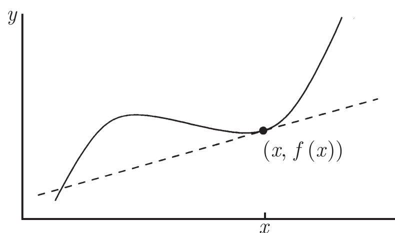


图　5-11

也可能在一个图像上给定的一点没有切线. 例如, 考虑 $y = |x|$ 的图像, 如图 5-12 所示. 该图像通过点 $(0,0)$, 但过那一点没有切线. 毕竟, 怎么可能会有切线? 不管怎么画, 都不能在那里同时顾及两边的图像, 因为它在原点处有一个尖点. 在后面的 5.2.10 节将返回到该例子.

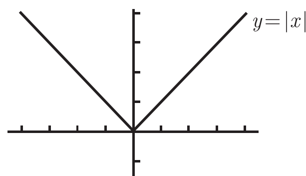


图　5-12

即使通过 $(x, f(x))$ 的切线存在, 你又该如何找到它呢? 回想一下, 为了描述一条直线, 仅仅需要提供两个信息：直线上的一点和该直线的斜率. 然后, 就可以使用点斜式来求直线方程. 其实, 我们已经有了一个要素了：直线通过点 $(x, f(x))$. 现在, 只需要求出斜率. 为了求解, 我们将玩一个游戏, 类似于在上一节中求瞬时速度玩的那个.

我们由选取一个靠近于 x(在它的左边或右边) 的数 z 开始, 并在曲线上画出点 $(z, f(z))$. 现在, 画一条通过点 $(x, f(x))$ 和 $(z, f(z))$ 的直线, 如图 5-13 所示.

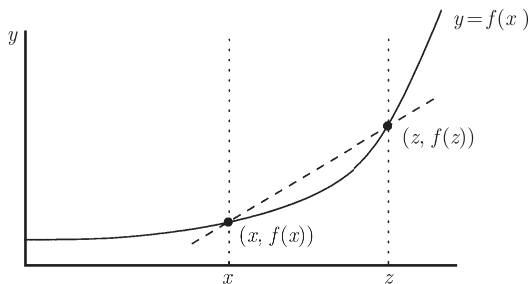


图　5-13

由于斜率是对边比邻边, 则虚线的斜率是

$$\frac{f(z)-f(x)}{z-x}.$$

现在, 当点 z 越来越接近 x, 但没有真正到达 x 的情况下, 以上直线的斜率应该变得越来越接近要找的切线的斜率. 因此, 显然有

$$\text{通过 }(x,f(x))\text{ 的切线的斜率} = \lim_{z\to x}\frac{f(z)-f(x)}{z-x}.$$

设 $h=z-x$, 可以看到, 当 $z\to x$ 时, 有 $h\to 0$, 从而也有

$$\text{通过 }(x,f(x))\text{ 的切线的斜率} = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}.$$

当然, 这只有当极限确实存在的时候才说得通!

5.2.6 导函数

在图 5-14 中, 我在曲线上画了通过三个不同的点的切线.

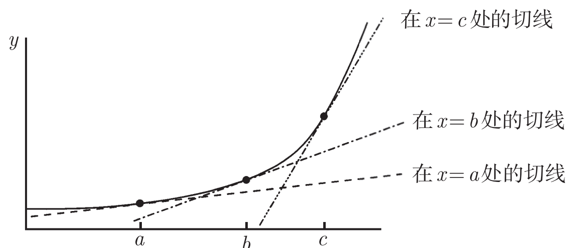


图 5-14

这些直线有不同的斜率. 也就是说, 切线的斜率取决于你选取的点 x 的值. 换句话说, 通过 $(x,f(x))$ 的切线的斜率是 x 的一个函数. 这个函数被称为 f 的**导数**, 并写作 f'. 我们说, 对 f 关于变量 x **求导**得到函数 f'. 根据上一节结尾部分的公式, 如果极限存在的话, 有

$$\boxed{f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}.}$$

在这种情况下, f 在 x 点**可导**. 如果对于某个特定的 x, 极限不存在, 那么 x 的值就没有在导函数 f' 的定义域里, 即 f 在 x 点**不可导**. 有很多原因会导致极限不存在. 比如说, 那里有一个尖角, 就像前述 $y=|x|$ 的例子中那样. 从更基本的层次上说, 如果 x 没有在 f 的定义域中, 那么甚至不可能画出点 $(x,f(x))$, 更不用说在那里画一条切线了.

回忆一下 5.2.3 节中瞬时速度的定义吧：

$$\text{在时刻 } t \text{ 的瞬时速度} = \lim_{h\to 0}\frac{f(t+h)-f(t)}{h},$$

其中 $f(t)$ 是汽车在时刻 t 的位置. 等号右边的表达式和上述 $f'(x)$ 的定义一样, 只

是用 x 代替了 t! 这就是说, 如果 $v(t)$ 是在时刻 t 的瞬时速度, 那么 $v(t)=f'(t)$. 速度正是位置关于时间的导数.

来看一个关于求导的例子. 如果 $f(x)=x^2$, 那么 $f'(x)$ 是什么呢? 计算过程和 5.2.3 节结尾部分很相似:

$$\begin{aligned} f'(x) &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{h\to 0}\frac{(x+h)^2-x^2}{h} \\ &= \lim_{h\to 0}\frac{x^2+2xh+h^2-x^2}{h} = \lim_{h\to 0}\frac{2xh+h^2}{h} \\ &= \lim_{h\to 0}(2x+h) = 2x. \end{aligned}$$

因此, $f(x)=x^2$ 的导数由 $f'(x)=2x$ 给出. 这意味着, 抛物线 $y=x^2$ 在点 (x,x^2) 的切线的斜率就是 $2x$. 让我们画出该曲线和一些切线来检验一下, 如图 5-15 所示.

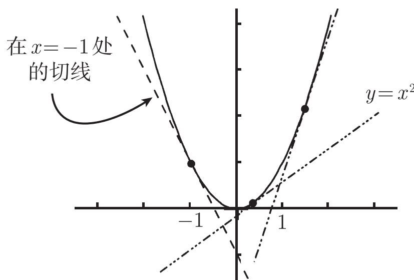


图　5-15

在 $x=-1$ 处的切线的斜率看起来的确是 -2, 这与公式 $f'(x)=2x$ 是一致的. (两倍的 -1 是 -2!) 其他切线也一样, 它们的斜率都是相应的 x 坐标的两倍.

5.2.7　作为极限比的导数

在导函数 $f'(x)$ 的公式中, 必须求出量 $f(x+h)$ 的值. 这个量是什么呢? 其实, 如果 $y=f(x)$, 将 x 变为 $x+h$, 那么 $f(x+h)$ 只是一个新的 y 值. 量 h 代表对 x 作了多少改变, 因此用量 Δx 作替换. 这里的符号 Δ 表示 "在 …… 中的变化", 因此 Δx 就是在 x 中的变化. (不要把 Δx 看作是 Δ 和 x 的乘积, 否则是错的!) 因此, 用 Δx 替换 h, 来重新写一下 $f'(x)$ 的公式:

$$f'(x)=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

好了, 情况是这样的. 由 (x,y) 开始, 其中 $y=f(x)$. 现在, 选取一个新的 x 值, 称之为 $x_{新}$. y 的值也会相应地变成 $y_{新}$, 这当然就是 $f(x_{新})$. 现在, 任意量的改变量正好是新值减去旧值, 因此有两个方程:

$$\Delta x = x_{新}-x \quad 和 \quad \Delta y = y_{新}-y.$$

第一个方程说的是 $x_{新}=x+\Delta x$, 因此第二个方程现在可以变形为

$$\Delta y = y_{新} - y = f(x_{新}) - f(x) = f(x + \Delta x) - f(x).$$

这就是上面 $f'(x)$ 定义中分数的分子! 这意味着

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

该公式的一个阐释是, x 中的一个小的变化产生了大约 $f'(x)$ 倍的 y 中的变化. 的确, 如果 $y = f(x) = x^2$, 那么在上一节已经看到 $f'(x) = 2x$. 让我们将精力集中在例如当 $x = 6$ 时的情况. 首先注意到, 由 $f'(x)$ 的公式可知 $f'(6) = 2 \times 6 = 12$. 因此, 如果取等式 $6^2 = 36$ 并将 6 作一点点改变, 36 将会变化 12 倍于此的量. 例如, 如果把 0.01 加到 6 上, 就应该将 0.12 加到 36 上. 因此, 我会猜 $(6.01)^2$ 应该差不多是 36.12. 事实上, 确切答案是 36.1201, 因此我的猜测确实接近.

那么, 为什么我没有得到确切答案呢? 原因是, $f'(x)$ 并不真正地等于 Δy 和 Δx 的比值, 它等于当 Δx 趋于 0 时该比值的极限. 这意味着, 如果没有离 6 太远的话, 可能会做得更好. 让我们来试着猜一下 $(6.0004)^2$ 的值吧. 将原始的 x 值 6 加上了 0.0004, 因此, y 值应该有 12 倍于此的改变, 也就是 0.0048. 因此, 我们猜测 $(6.0004)^2$ 大约是 36.0048. 这还不错 —— 真正的答案是 36.004 800 16, 两个数已经非常接近了! 对 6 的改变越小, 我们的方法计算出的结果就会越好.

当然, 魔力数字 12 仅仅当从 $x = 6$ 开始的时候才会起作用. 如果从 $x = 13$ 开始的话, 魔力数字就是 $f'(13)$ 了, 它就等于 $2 \times 13 = 26$. 因此, 我们知道 $13^2 = 169$, 那 $(13.0002)^2$ 是什么呢? 为了从 13 得到 13.0002, 必须加上 0.0002. 由于魔力数字是 26, 必须将 26 倍的 0.0002 加到 169 上来得到我们的猜测. 这就是说, 将 0.0052 加到 169 上并得出猜测结果是 169.0052. 再一次地, 这相当不错：$(13.0002)^2$ 实际上 169.005 200 04.

不管怎样, 我们在第 13 章讲解线性化时将返回到这些基本思想上来. 现在再来看看公式

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

等号右边的表达式是, 当 x 中的变化非常小时, y 中的变化与 x 中的变化的比值的极限. 假设 x 小得以至于其中的变化几乎注意不到. 现在我们不写 Δx, 它表示 “x 中的变化”, 而是写 $\mathrm{d}x$, 它表示 “x 中的十分微小的变化”. 对 y 也有类似的表示方法. 不幸的是, $\mathrm{d}x$ 和 $\mathrm{d}y$ 本身没有什么意义[①]; 尽管如此, 这给了我们灵感, 可以用一种不同的且更方便的方法来写导数：

$$\text{如果 } y = f(x), \text{ 那么可以用 } \frac{\mathrm{d}y}{\mathrm{d}x} \text{ 来代替 } f'(x).$$

例如, 如果 $y = x^2$, 那么 $\frac{\mathrm{d}y}{\mathrm{d}x} = 2x$. 事实上, 如果用 x^2 代替 y, 会得到对于一件事情

① “无穷小” 也有其理论, 但它超出了本书的范围!

的很多不同的表达方式：

$$f'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}(x^2)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}(x^2) = 2x.$$

作为另一个例子, 在 5.2.3 节, 我们看到过, 如果汽车在时刻 t 的位置是 $f(t) = 15t^2 + 7$, 那么它的速度是 $30t$. 回想一下, 速度就是 $f'(t)$, 这意味着 $f'(t) = 30t$. 但如果我们决定把位置称为 p, 从而 $p = 15t^2 + 7$, 便可以写 $\frac{\mathrm{d}p}{\mathrm{d}t} = 30t$. 这里的要点是, 不是所有的量都用 x 和 y 来表达, 你必须能够应对其他的字母.

总而言之, 量 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 是 y 关于 x 的导数. 如果 $y = f(x)$, 那么 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 和 $f'(x)$ 是一回事. 最后, 请记住, 量 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 实际上根本不是一个分数, 它是当 $\Delta x \to 0$ 时分数 $\frac{\Delta y}{\Delta x}$ 的极限.

5.2.8　线性函数的导数

让我们暂停一下喘口气, 回到一个简单的例子：假设 f 是线性的. 这意味着, 对于某个 m 和 b, $f(x) = mx + b$. 你认为 $f'(x)$ 会是什么？回想一下, 它度量的是, 曲线 $y = f(x)$ 在点 $(x, f(x))$ 处的切线的斜率. 在这个例子中, $y = mx + b$ 的图像就是斜率为 m、y 轴截距为 b 的一条直线. 显而易见, 该条直线上任意一点的切线就是这条直线本身！这意味着, 不管 x 取何值, $f'(x)$ 的值就应该是 m, 因为曲线 $y = mx + b$ 有固定的斜率 m. 用公式检验一下：

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(m(x+h)+b) - (mx+b)}{h} \\ &= \lim_{h \to 0} \frac{mh}{h} = \lim_{h \to 0} m = m. \end{aligned}$$

因此, 不管 x 取何值, $f'(x) = m$. 这就是说, 线性函数的导数是常数. 如你所想, 只有线性函数有固定的斜率 (这是所谓的中值定理的结果, 具体参见 11.3.1 节). 顺便说一下, 如果 f 是常数函数, 即 $f(x) = b$, 那么其斜率总是 0. 特别是, 对于所有的 x, $f'(x) = 0$. 因此, 这证明了常数函数的导数恒为 0.

5.2.9　二阶导数和更高阶导数

由于可以由一个函数 f 出发, 取其导数得到一个新的函数 f', 实际上可以采用这个新的函数, 再次求导. 最终得到导数的导数, 这被称为**二阶导**, 写作 f''.

例如, 如果 $f(x) = x^2$, 那么其导数为 $f'(x) = 2x$. 现在, 我们想要对此结果求导. 设 $g(x) = 2x$, 并试着求出 $g'(x)$. 由于 g 是一个线性函数, 其斜率为 2, 从上一节我们知道 $g'(x) = 2$. 因此, f 导数的导数是常数函数 2, 这样就证明了, 对于所有的 x, $f''(x) = 2$.

如果 $y = f(x)$, 那么我们已经看到, 可以用 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 代替 $f'(x)$. 对于二阶导有一种

相似的记号：

$$\text{如果 } y=f(x)\text{, 那么可以用 } \frac{\mathrm{d}^2y}{\mathrm{d}x^2} \text{ 代替 } f''(x).$$

在上述例子中, 如果 $y=f(x)=x^2$, 我们已经看到

$$f''(x)=\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}^2(x^2)}{\mathrm{d}x^2}=\frac{\mathrm{d}^2}{\mathrm{d}x^2}(x^2)=2.$$

这些都是对 $f(x)=x^2$ (关于 x) 的二阶导是常数函数 2 的有效的表达方式.

为什么要止步于求二阶导呢? 函数 f 的三阶导是 f 的导数的导数的导数. 这可是一长串 “导数”!, 你应该把 f 的三阶导看成是 f 二阶导的导数, 并且可以用以下任意一种方式写出：

$$f'''(x),\quad f^{(3)}(x),\quad \frac{\mathrm{d}^3y}{\mathrm{d}x^3}\quad \text{或}\quad \frac{\mathrm{d}^3}{\mathrm{d}x^3}(y).$$

记号 $f^{(3)}(x)$ 对于高阶导数尤其方便, 因为写那么多的撇号简直太傻了. 因此, 四阶导, 即三阶导的导数, 就可以写作 $f^{(4)}(x)$ 而不是 $f''''(x)$. 尽管如此, 对于低阶导数, 有时候用这种方式表示也会很方便, 比如将二阶导写成 $f^{(2)}(x)$ 而不是 $f''(x)$. 甚至也可能将一阶导写成 $f^{(1)}(x)$ 而不是 $f'(x)$, 因为只取了一次导数, 此外, 还可以用 $f^{(0)}(x)$ 代替 $f(x)$ 本身 (没有取导数!). 用这种方式, 任何导数都可以写成 $f^{(n)}(x)$ 的形式, 其中 n 为正整数.

5.2.10 何时导数不存在

在 5.2.5 节, 我提到过 $f(x)=|x|$ 的图像在原点处有一个尖点. 而这应该意味着, 在 $x=0$ 处导数不存在. 现在来看看为什么会是这样. 使用导数公式, 有

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\lim_{h\to 0}\frac{|x+h|-|x|}{h}.$$

我们感兴趣的是 $x=0$ 时会发生什么, 因此在以上等式链中用 0 替换 x, 得到

$$f'(0)=\lim_{h\to 0}\frac{f(0+h)-f(0)}{h}=\lim_{h\to 0}\frac{|0+h|-|0|}{h}=\lim_{h\to 0}\frac{|h|}{h}.$$

我们之前看到过这个极限! 事实上, 在 4.6 节, 该极限不存在. 这意味着, $f'(0)$ 的值无定义, 即 0 没有在 f' 的定义域中. 然而我们也看到过, 如果将它由一个双侧极限改为单侧极限, 那么以上极限存在. 特别是, 右极限是 1, 左极限是 -1. 这激发了**右导数**和**左导数**的思想, 其定义分别为

$$\lim_{h\to 0^+}\frac{f(x+h)-f(x)}{h}\quad \text{和}\quad \lim_{h\to 0^-}\frac{f(x+h)-f(x)}{h}.$$

它们看起来和普通导数的定义很相似, 只是双侧极限 (即当 $h\to 0$) 分别由右极限和左极限所代替. 跟在极限的情况一样, 如果左导数和右导数存在且相等, 那么实际的导数存在且有相同的值. 同时, 如果导数存在, 那么左右导数都存在且都等于

导数值.

不管怎样, 这里的要点是, 如果 $f(x) = |x|$, 那么在 $x = 0$ 处其右导数为 1, 左导数为 -1. 你相信吗? 让我们再来看看图 5-16. 当从原点出发沿着该曲线向右移动时, 它的斜率确实是 1 (事实上, 斜率始终为 1, 即如果 $x > 0$, $f'(x) = 1$). 类似地, 从原点出发沿着该曲线向左移动时, 它的斜率是 -1 (事实上, 如果 $x < 0$, $f'(x) = -1$). 由于左侧斜率不等于右侧斜率, 所以在 $x = 0$ 处导数不存在.

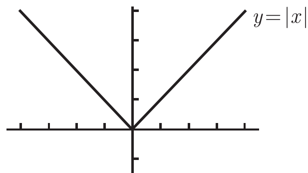


图　5-16

现在, 我们有了在其定义域内不是处处可导的连续函数. 很明显, 除了一个小点外, 它仍然是可导的. 事实上, 你可以有这样一个连续函数, 它是如此起伏多刺以至于它实际上在每一个单点 x 上都有一个尖角, 因此它在任意点上都不可导! 这种怪异的函数超出了本书的研究范围, 但我要顺便提及, 这种类型的函数可以用来为股价建模 —— 如果你曾经看到过股价的图像, 就会知道我说的起伏多刺是什么意思了. 不管怎样, 这里我的要点的是, 存在不可导的连续函数. 那么会有不连续的可导函数吗? 回答是否定的, 我们马上就会看到原因.

5.2.11　可导性和连续性

现在是时候将本章的两个重要概念联系在一起了. 我将要表明, 每一个可导函数也是连续的. 换言之, 如果你知道一个函数是可导的, 那么你将买一赠一, 获知该函数的连续性. 更确切地说, 我将要表明：

如果一个函数 f 在 x 上可导, 那么它在 x 上连续.

例如, 将在第 7 章证明, $\sin(x)$ 作为 x 的函数是可导的. 这将自动暗示它在 x 处也是连续的. 同样的结论也适用于其他的三角函数、指数函数和对数函数 (除了在它们的垂直渐近线处).

那么该如何证明我们这个重大断言呢? 先来看看我们想证明的是什么. 要证明 f 在 x 上连续, 需要证明

$$\lim_{u \to x} f(u) = f(x),$$

并且根据 5.1.1 节, 只有当等号两边同时存在时, 上式才成立! 在继续证明之前, 我想用 $h = u - x$ 作替换, 正如我们之前做过的. 在这种情况下, $u = x + h$, 并且当 $u \to x$ 时, 我们看到 $h \to 0$. 因此, 上式变为

$$\lim_{h \to 0} f(x + h) = f(x).$$

我们需要证明等号两边都存在且相等 —— 那样的话, 就完成任务了.

目标已经明确, 现在就让我们从实际知道的开始吧. 我们知道 f 在 x 上可导;

这意味着, $f'(x)$ 存在, 因此根据 f' 的定义, 极限

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

存在. 首先注意到, 上式中包含了 $f(x)$, 那么它一定存在, 否则上式就无从谈起. 因此, 我们已经有所进展：$f(x)$ 存在. 但我们仍然需要想些聪明的办法. 这里的技巧是, 由另一个极限开始：

$$\lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\times h\right).$$

一方面, 通过将它分成两个因子, 可以求出该极限为

$$\lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\times h\right)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\times\lim_{h\to 0}h=f'(x)\times 0=0.$$

由于所有涉及的极限都存在, 所以这样做没问题. (这里需要用到事实, $f'(x)$ 存在, 不然就有问题了.) 另一方面, 可以取原始极限并消去因子 h 得到

$$\lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\times h\right)=\lim_{h\to 0}(f(x+h)-f(x)).$$

比较一下这两个式子, 就会得到

$$\lim_{h\to 0}(f(x+h)-f(x))=0.$$

当然, $f(x)$ 的值根本不依赖于极限, 因此可以将它提出来, 得到

$$\left(\lim_{h\to 0}f(x+h)\right)-f(x)=0.$$

现在, 只需将 $f(x)$ 加到等号两边, 得到

$$\lim_{h\to 0}f(x+h)=f(x),$$

而这正是我们想要的! 特别是, 等号左边的极限存在并且等式成立. 因此, 我们证明了一个很好的结论：**可导函数必连续**. 不过要记住, 连续函数并不总是可导的!

第 6 章　求解微分问题

现在, 我们要看看如何应用上一章中的一些定理来求解微分问题. 我们可以利用公式求导, 但这很笨拙. 因此, 我们会看到一些能让生活变轻松的法则. 总之, 以下是我们在本章要讲解的内容:

- 使用定义求导;
- 使用乘积法则、商法则和链式求导法则;
- 求切线方程;
- 速度和加速度;
- 求导数伪装的极限;
- 如何对分段函数求导;
- 使用一个函数图像来画出其导函数的图像.

6.1　使用定义求导



假设要对 $f(x) = 1/x$ 关于 x 求导. 从上一章中可知, 导数的定义是

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h},$$

因此, 现在有

$$f'(x) = \lim_{h\to 0} \frac{\dfrac{1}{x+h} - \dfrac{1}{x}}{h}.$$

在分式中, 如果只是用 0 替换 h, 结果就会得到一个 $\dfrac{0}{0}$ 的不定式. 因此, 需要多计算一点. 在这里, 基本思想是通过通分来化简分子. 你会得到

$$f'(x) = \lim_{h\to 0} \frac{\dfrac{x-(x+h)}{x(x+h)}}{h} = \lim_{h\to 0} \frac{-h}{hx(x+h)}.$$

现在从分子分母中消去 h, 然后通过设 $h = 0$ 求极限值:

$$f'(x) = \lim_{h\to 0} \frac{-1}{x(x+h)} = \frac{-1}{x(x)} = -\frac{1}{x^2}.$$

也就是说,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x}\right) = -\frac{1}{x^2}.}$$

另一方面, 为了求 $f(x) = \sqrt{x}$ 的导数, 必须利用在 4.2 节中使用过的技巧. 具体如下:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{\sqrt{x+h} - \sqrt{x}}{h},$$

我们再次遇到 $\dfrac{0}{0}$ 的情况. 将分子和分母同时乘以分子的共轭表达式, 得到

$$f'(x) = \lim_{h\to 0} \frac{\sqrt{x+h} - \sqrt{x}}{h} \times \frac{\sqrt{x+h} + \sqrt{x}}{\sqrt{x+h} + \sqrt{x}} = \lim_{h\to 0} \frac{(x+h) - x}{h(\sqrt{x+h} + \sqrt{x})};$$

现在, 可以在分子上消去 x 这一项, 从分子和分母中消去 h, 然后求极限, 得到

$$f'(x) = \lim_{h\to 0} \frac{h}{h(\sqrt{x+h} + \sqrt{x})} = \lim_{h\to 0} \frac{1}{\sqrt{x+h} + \sqrt{x}} = \frac{1}{\sqrt{x} + \sqrt{x}} = \frac{1}{2\sqrt{x}}.$$

总而言之, 这就证明了

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}(\sqrt{x}) = \frac{1}{2\sqrt{x}}.}$$

现在, **使用导数的定义**, 你会如何求 $f(x) = \sqrt{x} + x^2$ 的导数呢? 即使你能够直接写出答案, 但我要求的是使用导数的定义, 所以你必须撇开一切诱惑并使用公式

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{(\sqrt{x+h} + (x+h)^2) - (\sqrt{x} + x^2)}{h}.$$

这看起来很杂乱, 但如果将它分成含有平方根的项和含有平方的项, 会看到

$$f'(x) = \lim_{h\to 0} \frac{\sqrt{x+h} - \sqrt{x}}{h} + \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h}.$$

我们知道该如何来求这两个极限; 刚刚第一个极限是 $1/2\sqrt{x}$, 而在 5.2.6 节求得第二个极限是 $2x$. 你应该试着不看前面的求解过程自己做一遍, 并确保得到正确答案

$$f'(x) = \frac{1}{2\sqrt{x}} + 2x.$$

现在是时候对 x^n 关于 x 求导了, 其中 n 是某个正整数. 设 $f(x) = x^n$, 那么有

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h}.$$

我们必须想办法处理 $(x+h)^n$. 有很多方法能处理该问题. 尝试最直接的方法, 那就是写出

$$(x+h)^n = (x+h)(x+h)\cdots(x+h).$$

在以上乘积中有 n 个因子. 如果将它们都乘开会很混乱, 但事实上, 不需要全部展开, 只需要开头部分. 如果从每一个因子中提取项 x, 将会有 n 个 x, 因而会在乘积中得到 x^n 这一项. 那是得到所有 x 因子的唯一方法, 因此有

$$(x+h)^n = (x+h)(x+h)\cdots(x+h) = x^n + \text{ 含有 } h \text{ 的项}$$

然而, 还需要再多做一点. 要是从第一个因子中提取 h, 然后从其他因子中提取 x, 又会怎样呢? 那样就会有一个 h 和 $(n-1)$ 个 x, 因此当将它们都乘起来的时候, 会得到 hx^{n-1}. 还有其他的方法来选择一个 h 和其余的 x (可以从第二个因子里提取 h, 然后从其他因子中提取 x; 或者, 从第三个因子里提取 h, 然后从其他因子中提取 x, 如此等等). 事实上, 有 n 种方法来选取一个 h 和其余的 x, 因此实际上有 n 个 hx^{n-1}. 加在一起, 会得到 nhx^{n-1}. 在展开式中, 其余每一项至少有两个 h, 因此，其余每一项就含有一个带 h^2 的因子. 总之, 可以写成

$$(x+h)^n = (x+h)(x+h)\cdots(x+h) = x^n + nhx^{n-1} + \text{ 含有因子 } h^2 \text{ 的项}$$

稍作整理：将用 $h^2\times$(垃圾) 代表 "含有因子 h^2 的项", 其中 "垃圾" 就是含有 x 和 h 的多项式. 也就是说,

$$(x+h)^n = (x+h)(x+h)\cdots(x+h) = x^n + nhx^{n-1} + h^2\times(\text{垃圾}).$$

现在, 可以将以上形式带入导数的公式里：

$$f'(x) = \lim_{h\to 0}\frac{(x+h)^n - x^n}{h} = \lim_{h\to 0}\frac{x^n + nhx^{n-1} + h^2\times\ (\text{垃圾}) - x^n}{h}.$$

x^n 这一项被消去了, 然后可以分子分母消去 h：

$$f'(x) = \lim_{h\to 0}\frac{nhx^{n-1} + h^2\times(\text{垃圾})}{h} = \lim_{h\to 0}(nx^{n-1} + h\times(\text{垃圾})).$$

当 $h\to 0$ 时, 第二项趋于 0, 而第一项仍然是 nx^{n-1}. 因此, 我们得出结论, 当 n 是一个正整数时,

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^n) = nx^{n-1}.$$

事实上, 我们将会在 9.5.1 节中证明, 当 a 是任意实数时,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}(x^a) = ax^{a-1}.}$$

用文字表述就是：提取次数, 将它放在最前面作系数, 然后再将次数减少 1.

再来好好看看以上公式. 首先, 当 $a=0$ 时, x^a 是常数函数 1. 其导数是 $0x^{-1}$, 结果就是 0. 这和 5.2.8 节中的计算一致. 总而言之,

$$\text{如果 } C \text{ 是常数, 那么 } \frac{\mathrm{d}}{\mathrm{d}x}(C)=0.$$

现在, 如果 $a=1$, 那么 x^a 就是 x. 根据公式, 其导数为 $1x^0$, 也就是常数函数 1. 同样, 这和 5.2.8 节中的结果一致. 因此, 可以确认

$$\frac{\mathrm{d}}{\mathrm{d}x}(x)=1.$$

当 $a=2$ 时, 可以看到 x^2 关于 x 的导数是 $2x^1$, 也就是 $2x$. 这和之前的结论一致. 类似地, 当 $a=-1$ 时, 可以使用公式并看到 x^{-1} 的导数是 $-1\times x^{-2}$. 事实上, 这就是说 $1/x$ 的导数是 $-1/x^2$, 这一点我们在本节开始的时候已经知道了! 这个例子会经常出现, 你应该特别掌握.

现在, 来尝试一些指数为分数的情况. 当 $a=\dfrac{1}{2}$ 时, $x^{1/2}$ 关于 x 的导数是 $\dfrac{1}{2}x^{-1/2}$. 根据指数法则 (关于这些的回顾请参见 9.1.1 节), 可以重写并看到 $\sqrt{x}$ 的导数是 $1/2\sqrt{x}$, 这正是之前所求得的结果. 再次地, 它也会经常出现, 所以要特别掌握, 以避免将指数 $\dfrac{1}{2}$ 和 $-\dfrac{1}{2}$ 搞错了. 最后看一下 $a=\dfrac{1}{3}$ 的情况. 公式告诉我们

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^{1/3})=\frac{1}{3}x^{1/3-1}=\frac{1}{3}x^{-2/3}.$$

使用指数法则 (再次地, 你可以在 9.1.1 节中找到它), 将其重写成

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sqrt[3]{x})=\frac{1}{3\sqrt[3]{x^2}}.$$

这个稍微复杂一些, 所以你不必费心去记, 只需能够使用上述 x^a 关于 x 的求导公式来推导出它就可以了.

6.2 用更好的办法求导

所有这些折腾极限的求导不免有些烦琐乏味. 幸运的是, 一旦你做完了它们, 就可以根据一些简单的法则由已经求得的导数来构造其他的导数了. 让我们定义一个函数

$$f(x)=\frac{3x^7+x^4\sqrt{2x^5+15x^{4/3}-23x+9}}{6x^2-4}.$$

对类似这样一个函数求导的关键是, 理解它是如何由简单函数合成的. 在 6.2.6 节, 我们将会看到如何使用简单的运算 (函数的常数倍、函数的加法、减法、乘法、除法以及复合函数) 用形如 x^a 的原子来构造 f, 而对于 x^a 我们已经知道如何求导了. 首先, 需要看看求导将如何受到这些运算的影响; 然后, 再回来求以上那个难以处理的函数 f 的 $f'(x)$. (以下法则的正式证明参见附录 A 中的 A.6 节, 而在 6.2.7

节中会有对其中一些法则的直观证明.)

6.2.1　函数的常数倍

处理一个函数的常数倍很容易：只需在求导后，用常数乘以该函数的导数就可以了．例如，我们知道 x^2 的导数是 $2x$，因此 $7x^2$ 的导数就是 7 倍的 $2x$，即 $14x$. $-x^2$ 的导数是 $-2x$，因为你可以认为前面的负号是用 -1 做乘法的结果. 事实上，有一个简单的方法来求 x^a 的常数倍的导数：将指数拖下来，用它和常数相乘，然后将指数降低一次．因此，对于 $7x^2$ 的导数，将 2 拖下来，用它和 7 相乘得到系数 14，然后将 x 指数降低一次得到 $14x^1$，也就是 $14x$．类似地，为了求 $13x^4$ 的导数，用 4 乘以 13，得到系数为 52，然后将 x 指数降低一次得到 $52x^3$.

6.2.2　函数和与函数差

对函数和与函数差求导则更容易：对每一部分求导，然后再相加或相减就可以了．例如，

$$3x^5 - 2x^2 + \frac{7}{\sqrt{x}} + 2$$

关于 x 的导数是什么呢？首先，将 $1/\sqrt{x}$ 写成 $x^{-1/2}$，这意味着，必须要对 $3x^5 - 2x^2 + 7x^{-1/2} + 2$ 求导．使用刚刚看到的常数倍的求导方法，$3x^5$ 的导数是 $15x^4$．类似地，$-2x^2$ 的导数是 $-4x$，$7x^{-1/2}$ 的导数是 $-\frac{7}{2}x^{-3/2}$．最后，2 的导数是 0，因为 2 是一个常数．也就是说，只要是求导，在结尾的 $+2$ 就是无关紧要的．因此，将这些值组合在一起，得到

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(3x^5 - 2x^2 + \frac{7}{\sqrt{x}} + 2\right) = \frac{\mathrm{d}}{\mathrm{d}x}(3x^5 - 2x^2 + 7x^{-1/2} + 2) = 15x^4 - 4x - \frac{7}{2}x^{-3/2}.$$

顺便说一下，如果意识到可以将 $x^{3/2}$ 写成 $x\sqrt{x}$，也可以将以上导数写作

$$15x^4 - 4x - \frac{7}{2}\frac{1}{x\sqrt{x}}.$$

类似地，$x^{5/2}$ 就是 $x^2\sqrt{x}$，$x^{7/2}$ 就是 $x^3\sqrt{x}$，等等.

6.2.3　通过乘积法则求积函数的导数

处理函数乘积的时候要更麻烦些 —— 不能只是将两个导数乘在一起．例如，不做展开 (那样太费时间了)，我们想要求

$$h(x) = (x^5 + 2x - 1)(3x^8 - 2x^7 - x^4 - 3x)$$

的导数．设 $f(x) = x^5 + 2x - 1$ 及 $g(x) = 3x^8 - 2x^7 - x^4 - 3x$．函数 h 是 f 和 g 的乘积．我们可以很容易地写出 f 和 g 的导数，它们是 $f'(x) = 5x^4 + 2$ 及 $g'(x) = 24x^7 - 14x^6 - 4x^3 - 3$．如前所述，简单认为乘积 h 的导数是这两个导数的乘积是不正确的．也就是说，$h'(x) \neq \left(5x^4 + 2\right)\left(24x^7 - 14x^6 - 4x^3 - 3\right)$．当然，说

$h'(x)$ 不是什么是没有用的, 需要说它是什么!

事实上, 需要混合搭配. 也就是说, 取 f 的导数并用它和 g 相乘 (不是 g 的导数). 然后, 也需要取 g 的导数并用它和 f 相乘. 最后, 将它们加在一起. 具体如下：

乘积法则 (版本 1) 如果 $h(x) = f(x)g(x)$, 那么 $h'(x) = f'(x)g(x) + f(x)g'(x)$.

因此, 对于例子中的 $h(x) = \left(x^5 + 2x - 1\right)\left(3x^8 - 2x^7 - x^4 - 3x\right)$, 我们将 h 写成 f 和 g 的乘积并分别求它们的导数. 将结果汇总一下, 取每一列分别对应 f 和 g：

$$\begin{aligned} f(x) &= x^5 + 2x - 1 & g(x) &= 3x^8 - 2x^7 - x^4 - 3x \\ f'(x) &= 5x^4 + 2 & g'(x) &= 24x^7 - 14x^6 - 4x^3 - 3. \end{aligned}$$

现在, 可以使用乘积法则并做一些交叉相乘. 你看, 需要用左下方的 $f'(x)$ 和右上方的 $g(x)$ 相乘, 然后用左上方的 $f(x)$ 和右下方的 $g'(x)$ 相乘, 并将它们相加在一起. 这样得到

$$\begin{aligned} h'(x) &= f'(x)g(x) + f(x)g'(x) \\ &= (5x^4 + 2)(3x^8 - 2x^7 - x^4 - 3x) \\ &\quad + (x^5 + 2x - 1)(24x^7 - 14x^6 - 4x^3 - 3). \end{aligned}$$

可以将这个结果乘开, 但这会比将原始函数 h 乘开然后求导还要糟. 就让它这样吧.

还有另外一种方式来写乘积法则. 确实有时候, 必须处理 $y =$ 用 x 表示的项, 而不是 $f(x)$ 的形式. 例如, 假设 $y = \left(x^3 + 2x\right)(3x + \sqrt{x} + 1)$, $\mathrm{d}y/\mathrm{d}x$ 是什么呢? 在这种情况下, 令 $u = \left(x^3 + 2x\right)$ 及 $v = (3x + \sqrt{x} + 1)$ 会更容易一些. 然后, 可以使用以上形式的乘积法则并作一些替换：首先, u 替换 $f(x)$, 这样就使 $\mathrm{d}u/\mathrm{d}x$ 替换 $f'(x)$; 对于 v 和 $g(x)$ 也做同样的操作. 于是得到

乘积法则 (版本 2) 如果 $y = uv$, 则

$$\frac{\mathrm{d}y}{\mathrm{d}x} = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x}.$$

因此, 在例子中有

$$\begin{aligned} u &= x^3 + 2x & v &= 3x + \sqrt{x} + 1 \\ \frac{\mathrm{d}u}{\mathrm{d}x} &= 3x^2 + 2 & \frac{\mathrm{d}v}{\mathrm{d}x} &= 3 + \frac{1}{2\sqrt{x}}. \end{aligned}$$

这意味着

$$\frac{\mathrm{d}y}{\mathrm{d}x} = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x} = (3x + \sqrt{x} + 1)(3x^2 + 2) + (x^3 + 2x)\left(3 + \frac{1}{2\sqrt{x}}\right).$$

现在, 要是你有一个三项的乘积又会怎样呢? 例如, 假设

$$y = (x^2 + 1)(x^2 + 3x)(x^5 + 2x^4 + 7),$$

而你想要求 $\mathrm{d}y/\mathrm{d}x$. 可以将它乘开再求导, 或者使用适用于三项的乘积法则：

> **乘积法则 (三个变量)**　如果 $y = uvw$, 那么
> $$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}x}vw + u\frac{\mathrm{d}v}{\mathrm{d}x}w + uv\frac{\mathrm{d}w}{\mathrm{d}x}.$$

在解答例子之前, 先来看一个记住以上公式的小窍门：把 uvw 加三次, 但对于每一项, 要将 $\mathrm{d}/\mathrm{d}x$ 放在不同的变量之前. (同样的诀窍适用于四个或更多个变量 —— 每一个变量都要进行一次微分运算!) 不管怎样, 在例子中, 要令 $u = x^2+1$, $v = x^2+3x$, $w = x^5 + 2x^4 + 7$, 这样, y 就是乘积 uvw. 我们有 $\mathrm{d}u/\mathrm{d}x = 2x$, $\mathrm{d}v/\mathrm{d}x = 2x + 3$, $\mathrm{d}w/\mathrm{d}x = 5x^4 + 8x^3$. 根据以上公式, 有

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\mathrm{d}u}{\mathrm{d}x}vw + u\frac{\mathrm{d}v}{\mathrm{d}x}w + uv\frac{\mathrm{d}w}{\mathrm{d}x} \\ &= (2x)(x^2+3x)(x^5+2x^4+7) + (x^2+1)(2x+3)(x^5+2x^4+7) \\ &\quad +(x^2+1)(x^2+3x)(5x^4+8x^3).\end{aligned}$$

由于没有将以上 y 的原始表达式展开并化简, 显然我也不会化简这个导数! 不过, 我确实要提醒的是, 你不总是能将所有的一切都展开. 有时候只能使用乘积法则. 例如, 当你在下一章学了如何对三角函数求导之后, 只能使用乘积法则来求像 $x\sin(x)$ 这样的导数. 但真的不能将这个表达式展开 —— 它已经是展开的形式了. 因此, 如果想要对它关于 x 求导, 那就避免不了要使用乘积法则.

6.2.4　通过商法则求商函数的导数

处理商的方式与处理乘积的方式类似, 只是法则稍有不同. 假设想对

$$h(x) = \frac{2x^3 - 3x + 1}{x^5 - 8x^3 + 2}$$

关于 x 求导. 可以令 $f(x) = 2x^3 - 3x + 1$ 及 $g(x) = x^5 - 8x^3 + 2$, 然后将 h 写成 f 和 g 的商, 或 $h(x) = f(x)/g(x)$. 以下就是商法则：

> **商法则 (版本 1)**　如果 $h(x) = \dfrac{f(x)}{g(x)}$, 那么
> $$h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}.$$

注意到除了正号变成了负号外, 等号右边分式的分子与乘积法则中的分子是一样的. 在例子中, 需要对 f 和 g 求导并将结果汇总如下：

$$\begin{aligned} f(x) &= 2x^3 - 3x + 1 & g(x) &= x^5 - 8x^3 + 2 \\ f'(x) &= 6x^2 - 3 & g'(x) &= 5x^4 - 24x^2. \end{aligned}$$

根据商法则, 由于 $h(x) = f(x)/g(x)$, 有

$$\begin{aligned} h'(x) &= \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2} \\ &= \frac{(6x^2-3)(x^5-8x^3+2)-(2x^3-3x+1)(5x^4-24x^2)}{(x^5-8x^3+2)^2}. \end{aligned}$$

跟乘积法则一样, 这里还有另外一种版本. 如果你面对

$$y = \frac{3x^2+1}{2x^8-7},$$

并想求出 $\mathrm{d}y/\mathrm{d}x$, 那么就从设 $u = 3x^2+1$ 及 $v = 2x^8-7$ 开始, 这样 $y = u/v$. 现在我们使用:

> **商法则 (版本 2)** 如果 $y = \dfrac{u}{v}$, 那么
>
> $$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2}.$$

汇总表如下:

$$\begin{aligned} u &= 3x^2+1 & v &= 2x^8-7 \\ \frac{\mathrm{d}u}{\mathrm{d}x} &= 6x & \frac{\mathrm{d}v}{\mathrm{d}x} &= 16x^7. \end{aligned}$$

根据商法则,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2} = \frac{(2x^8-7)(6x)-(3x^2+1)(16x^7)}{(2x^8-7)^2}.$$

如你所见, 商的情况并不比乘积的难多少 (就是更杂乱了些).

6.2.5 通过链式求导法则求复合函数的导数

假设 $h(x) = \left(x^2+1\right)^{99}$, 你想要求 $h'(x)$. 将它展开来求乘积是很可笑的 (这样的话, 就必须用 x^2+1 和它本身相乘 99 次, 那将是很耗时的). 使用乘积法则也会很荒唐, 因为需要使用很多很多次.

相反, 将 h 看作是两个函数 f 和 g 的复合, 其中 $g(x) = x^2+1$, $f(x) = x^{99}$. 确实, 如果取一个 x, 将它放入 g 中, 会得到 x^2+1. 现在, 如果把它放入 f, 会得到 $\left(x^2+1\right)^{99}$, 即 $h(x)$. 这样, 就把 $h(x)$ 写成了 $f(g(x))$. (更多有关复合函数的内容请参见 1.3 节.) 现在, 可以应用**链式求导法则**了:

> **链式求导法则 (版本 1)** 如果 $h(x) = f(g(x))$, 那么 $h'(x) = f'(g(x))g'(x)$.

该公式看起来有点棘手. 让我们分解一下. 第二个因子很简单, 它正好是 g 的导数. 那么第一个因子呢? 好吧, 必须对 f 求导, 然后求其在 $g(x)$ 而不是 x 处的结果.

在例子中, 有 $f(x) = x^{99}$, 这样 $f'(x) = 99x^{98}$. 也有 $g(x) = x^2+1$, 故 $g'(x) =$

$2x$. 第二个因子就是 $2x$. 那么第一个因子呢？好, 取 $f'(x)$, 但现在不是将 x, 而是将 x^2+1 (因为这就是 $g(x)$) 放入其中. 也就是说, $f'(g(x)) = f'\left(x^2+1\right) = 99\left(x^2+1\right)^{98}$. 现在, 将这两个因子乘起来就会得到

$$h'(x) = f'(g(x))g'(x) = 99(x^2+1)^{98}(2x) = 198x(x^2+1)^{98}.$$

确实, 这看上去有点复杂. 还有另一种方法来求解这个问题.

我们由 $y = \left(x^2+1\right)^{99}$ 开始, 想要求 $\mathrm{d}y/\mathrm{d}x$. $\left(x^2+1\right)$ 这一项让问题变得复杂, 因此就称它为 u. 这意味着 $y = u^{99}$, 其中 $u = x^2+1$. 现在, 可以借助链式求导法则的另一个版本了：

链式求导法则 (版本 2) 如果 y 是 u 的函数, 并且 u 是 x 的函数, 那么

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}.$$

因此, 在我们的例子中, 有

$$y = u^{99} \qquad u = x^2+1$$
$$\frac{\mathrm{d}y}{\mathrm{d}u} = 99u^{98} \quad \frac{\mathrm{d}u}{\mathrm{d}x} = 2x.$$

使用以上框中的链式求导法则, 可以看到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = 99u^{98} \times 2x = 198xu^{98}.$$

现在, 只需用 x^2+1 替换 u, 便可得到 $\mathrm{d}y/\mathrm{d}x = 198x\left(x^2+1\right)^{98}$, 一如我们之前求出的.

下面是另一个简单的例子. 如果 $y = \sqrt{x^3-7x}$, 那么 $\mathrm{d}y/\mathrm{d}x$ 是什么呢？设 $u = x^3-7x$, 这样 $y = \sqrt{u}$. 汇总表如下：

$$y = \sqrt{u} \qquad u = x^3-7x$$
$$\frac{\mathrm{d}y}{\mathrm{d}u} = \frac{1}{2\sqrt{u}} \quad \frac{\mathrm{d}u}{\mathrm{d}x} = 3x^2-7.$$

因此, 根据链式求导法则有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{2\sqrt{u}} \times (3x^2-7) = \frac{3x^2-7}{2\sqrt{u}}.$$

现在, 只需在分母中替换掉 u. 由于 $u = x^3-7x$, 可以看到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{3x^2-7}{2\sqrt{x^3-7x}}.$$

一旦你掌握了, 事情就没有看上去那么难.

链式求导法则的两点简短说明. 首先, 为什么称它为链式求导法则呢？以 x 开始, 就会得到 u; 然后, 取 u 会得到 y. 这样通过额外的变量 u, 从 x 到 y 便形成了一种链. 其次, 你可能会认为链式求导法则是显而易见的. 毕竟, 在前面的方框公

式中, 不是能够消去因子 $\mathrm{d}u$ 吗? 回答是否定的. 回想一下, 诸如 $\mathrm{d}y/\mathrm{d}u$ 和 $\mathrm{d}u/\mathrm{d}x$ 这样的表达式其实不是分数, 它们是分数的极限 (更多详情参见 5.2.7 节). 不过幸好, 它们经常表现得就好像是分数 (显然这里它们就是如此).

链式求导法则实际上可以同时多次运用. 例如, 令

$$y = ((x^3 - 10x)^9 + 22)^8,$$

那么 $\mathrm{d}y/\mathrm{d}x$ 是什么呢? 我们令 $u = x^3 - 10x$ 及 $v = u^9 + 22$, 这样 $y = v^8$. 然后, 使用一个更长形式的链式求导法则:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}v}\frac{\mathrm{d}v}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}.$$

稍作思考, 你就不会弄错: y 是 v 的函数, v 是 u 的函数, u 是 x 的函数, 因此公式只可能是现在这个样子! 不管怎样, 我们有

$$y = v^8 \qquad v = u^9 + 22 \qquad u = x^3 - 10x$$

$$\frac{\mathrm{d}y}{\mathrm{d}v} = 8v^7 \quad \frac{\mathrm{d}v}{\mathrm{d}u} = 9u^8 \qquad \frac{\mathrm{d}u}{\mathrm{d}x} = 3x^2 - 10.$$

将所有的一切代入, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}v}\frac{\mathrm{d}v}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = (8v^7)(9u^8)(3x^2 - 10).$$

快大功告成了, 但还需要除掉 u 项和 v 项. 首先, 用 $u^9 + 22$ 替换 v:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = (8v^7)(9u^8)(3x^2 - 10) = (8(u^9 + 22)^7)(9u^8)(3x^2 - 10).$$

然后用 $x^3 - 10x$ 替换 u, 并合并因子 8 和 9, 得到真正的答案:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = (8(u^9 + 22)^7)(9u^8)(3x^2 - 10) = 72((x^3 - 10x)^9 + 22)^7(x^3 - 10x)^8(3x^2 - 10).$$

以上主要使用了链式求导法则的第二种形式, 但有时应用链式求导法则的第一种形式也会事半功倍. 例如, 对于某个函数 g 和 h, 如果知道 $h(x) = \sqrt{g(x)}$, 且 $g(5) = 4$ 以及 $g'(5) = 7$, 那么仍然可以求出 $h'(5)$. 我们设 $f(x) = \sqrt{x}$, 这样 $h(x) = f(g(x))$, 然后使用上述公式 $h'(x) = f'(g(x))g'(x)$. 由于 $f(x) = \sqrt{x}$, 有 $f'(x) = 1/2\sqrt{x}$; 因此,

$$h'(x) = f'(g(x))g'(x) = \frac{1}{2\sqrt{g(x)}}g'(x).$$

现在, 将 $x = 5$ 代入, 得到

$$h'(5) = \frac{1}{2\sqrt{g(5)}}g'(5).$$

由于 $g(5) = 4$ 及 $g'(5) = 7$, 有

$$h'(5) = \frac{1}{2\sqrt{4}}(7) = \frac{7}{4}.$$

再来看一个例子: 假设 $j(x) = g(\sqrt{x})$, 其中 g 的情况如上. $j'(25)$ 会是什么呢? 现

在, 有 $j(x) = g(f(x))$, 其中 $f(x) = \sqrt{x}$. 这一次的结果是

$$j'(x) = g'(f(x))f'(x) = g'(\sqrt{x})\frac{1}{2\sqrt{x}}.$$

因此, 如果 $x = 25$, 由于 $g'(5) = 7$, 有

$$j'(25) = g'(\sqrt{25})\frac{1}{2\sqrt{25}} = g'(5)\frac{1}{10} = \frac{7}{10}$$

比较一下这两个例子可知：复合的顺序非常重要!

6.2.6 那个难以处理的例子

回到开头提到的函数 f：

$$f(x) = \frac{3x^7 + x^4\sqrt{2x^5 + 15x^{4/3} - 23x + 9}}{6x^2 - 4}.$$

为了求出 $f'(x)$, 必须使用前几节中的法则将 f 分解为较简单的函数. 使用函数记号 (前述所有法则的第一种形式) 是一个不错的主意. 现在就试着做一下吧!

不过这里, 我将使用所有法则的第二种形式. 设 $y = f(x)$, 并试着求出 $\mathrm{d}y/\mathrm{d}x$. 首先, 注意到 y 是两部分的商：$u = 3x^7 + x^4\sqrt{2x^5 + 15x^{4/3} - 23x + 9}$ 及 $v = 6x^2 - 4$. 我们将使用商法则来处理这个分式, 因此需要 $\mathrm{d}u/\mathrm{d}x$ 和 $\mathrm{d}v/\mathrm{d}x$. 第二个非常好计算, 它就是 $12x$. 第一个有点难度. 让我们把目前已知的汇总一下：

$$u = 3x^7 + x^4\sqrt{2x^5 + 15x^{4/3} - 23x + 9} \qquad v = 6x^2 - 4$$
$$\frac{\mathrm{d}u}{\mathrm{d}x} = ??? \qquad \frac{\mathrm{d}v}{\mathrm{d}x} = 12x.$$

如果知道 $\mathrm{d}u/\mathrm{d}x$, 就可以使用商法则来完成运算. 因此, 要求出 $\mathrm{d}u/\mathrm{d}x$.

首先, 注意到 u 是 $q = 3x^7$ 和一个难解的量 r 的和, 其中 r 的定义为 $r = x^4\sqrt{2x^5 + 15x^{4/3} - 23x + 9}$. 我们需要这两部分的导数. q 的导数很简单, 它就是 $21x^6$. 现在, r 是 $w = x^4$ 和 $z = \sqrt{2x^5 + 15x^{4/3} - 23x + 9}$ 的乘积, 因此, 必须使用乘积法则来求 $\mathrm{d}r/\mathrm{d}x$. 目前有

$$w = x^4 \qquad z = \sqrt{2x^5 + 15x^{4/3} - 23x + 9}$$
$$\frac{\mathrm{d}w}{\mathrm{d}x} = 4x^3 \quad \frac{\mathrm{d}z}{\mathrm{d}x} = ???$$

真要命, 我们又不知道 $\mathrm{d}z/\mathrm{d}x$ 是什么, 所以需要求出它. 这里, 取一个大的表达式 (不妨称之为 t) 的平方根. 特别是, 如果 $t = 2x^5 + 15x^{4/3} - 23x + 9$, 那么 $z = \sqrt{t}$. 现在, 可以真正地求导了! 让我们给出最后一张汇总表：

$$t = 2x^5 + 15x^{4/3} - 23x + 9 \qquad z = \sqrt{t}$$
$$\frac{\mathrm{d}t}{\mathrm{d}x} = 10x^4 + 20x^{1/3} - 23 \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = \frac{1}{2\sqrt{t}}.$$

根据链式求导法则 (将变量改成我们需要的字母),

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{1}{2\sqrt{t}}(10x^4 + 20x^{1/3} - 23).$$

用 t 的定义 $2x^5 + 15x^{4/3} - 23x + 9$ 替换 t, 可以看到

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{10x^4 + 20x^{1/3} - 23}{2\sqrt{2x^5 + 15x^{4/3} - 23x + 9}}.$$

太棒了! 终于得到了 $\mathrm{d}z/\mathrm{d}x$. 现在可以将上表中的问号补充完整了：

$$\begin{array}{ll} w = x^4 & z = \sqrt{2x^5 + 15x^{4/3} - 23x + 9} \\ \dfrac{\mathrm{d}w}{\mathrm{d}x} = 4x^3 & \dfrac{\mathrm{d}z}{\mathrm{d}x} = \dfrac{10x^4 + 20x^{1/3} - 23}{2\sqrt{2x^5 + 15x^{4/3} - 23x + 9}}. \end{array}$$

现在, 再往前看：试图求出 $\mathrm{d}r/\mathrm{d}x$, 其中 $r = wz$. 让我们使用乘积法则：

$$\frac{\mathrm{d}r}{\mathrm{d}x} = z\frac{\mathrm{d}w}{\mathrm{d}x} + w\frac{\mathrm{d}z}{\mathrm{d}x}.$$

再一次地, 注意到对于变量你必须灵活处理, 它们不会总是 u 和 v! 不管怎样, 作相应替换得到

$$\frac{\mathrm{d}r}{\mathrm{d}x} = \left(\sqrt{2x^5 + 15x^{4/3} - 23x + 9}\right)(4x^3) + (x^4)\frac{10x^4 + 20x^{1/3} - 23}{2\sqrt{2x^5 + 15x^{4/3} - 23x + 9}}.$$

通分并化简上式, 得到 (自己检验一下!)

$$\frac{\mathrm{d}r}{\mathrm{d}x} = \frac{26x^8 + 140x^{13/3} - 207x^4 + 72x^3}{2\sqrt{2x^5 + 15x^{4/3} - 23x + 9}}.$$

现在返回到 u. 我们已经看到 $u = q + r$, 其中有 $q = 3x^7$, $r = x^4\sqrt{2x^5 + 15x^{4/3} - 23x + 9}$. 我们知道 $\mathrm{d}q/\mathrm{d}x = 21x^6$, 并且已经解出了杂乱的 $\mathrm{d}r/\mathrm{d}x$ 的公式, 因此只要把它们加在一起, 就会得到

$$\frac{\mathrm{d}u}{\mathrm{d}x} = 21x^6 + \frac{26x^8 + 140x^{13/3} - 207x^4 + 72x^3}{2\sqrt{2x^5 + 15x^{4/3} - 23x + 9}}.$$

最后, 返回到商和 $\dfrac{\mathrm{d}v}{\mathrm{d}x}$ 的计算, 补充进 $\mathrm{d}u/\mathrm{d}x$ 得到

$$\begin{array}{ll} u = 3x^7 + x^4\sqrt{2x^5 + 15x^{4/3} - 23x + 9} & v = 6x^2 - 4 \\ \dfrac{\mathrm{d}u}{\mathrm{d}x} = 21x^6 + \dfrac{26x^8 + 140x^{13/3} - 207x^4 + 72x^3}{2\sqrt{2x^5 + 15x^{4/3} - 23x + 9}} & \dfrac{\mathrm{d}v}{\mathrm{d}x} = 12x. \end{array}$$

由于 $y = u/v$, 只需使用标准的商法则

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2}.$$

拆分和消去之后, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{21x^6 + \dfrac{26x^8 + 140x^{13/3} - 207x^4 + 72x^3}{2\sqrt{2x^5 + 15x^{4/3} - 23x + 9}}}{6x^2 - 4}$$
$$- \frac{\left(3x^7 + x^4\sqrt{2x^5 + 15x^{4/3} - 23x + 9}\right)(12x)}{(6x^2 - 4)^2}.$$

终于完成解答了! 这个解答确实不美观, 但它无疑是有效的.

6.2.7　乘积法则和链式求导法则的理由

在附录 A 的 A.6.3 节和 A.6.5 节中, 可以找到乘积法则和链式求导法则的正式证明, 但先对为什么这些法则会起作用有个直观概念, 也是一个不错的主意. 因此, 让我们来快速地看一下吧.

就乘积法则来说, 将使用 6.2.3 节中该法则的第二种形式. 我们以两个量 u 和 v 开始, 它们都依赖于某个变量 x. 我们想知道, 如果 x 有一个小的变化量 Δx, 乘积 uv 将如何变化. 显然, u 会变成 $u + \Delta u$, v 会变成 $v + \Delta v$, 因此乘积变成了 $(u + \Delta u)(v + \Delta v)$. 可以通过想象一个边长分别为 u 和 v 个单位长度的矩形来理解. 该矩形的形状发生了一点变化, 其新的边长分别是 $u + \Delta u$ 和 $v + \Delta v$ 个单位长度, 如图 6-1 所示.

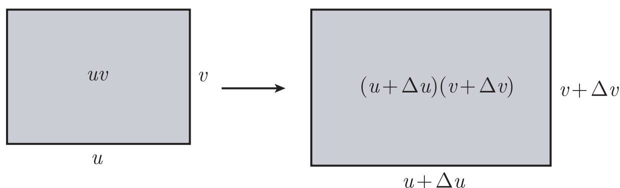


图　6-1

乘积 uv 和 $(u + \Delta u)(v + \Delta v)$ 正好分别是两个矩形的面积, 单位是平方单位. 那么面积有多大改变呢? 将这两个矩形重叠起来看一下, 如图 6-2 所示.

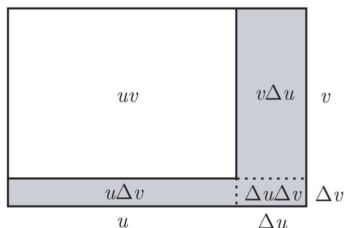


图　6-2

面积的差恰好是灰色 L 形区域的面积. 该区域由两个狭长的矩形 (面积分别为 $v\Delta u$ 和 $u\Delta v$ 平方单位) 以及一个小矩形 (面积为 $\Delta u\Delta v$ 平方单位) 组成. 由于面积的改变是 $\Delta(uv)$ 平方单位, 这就证明了

$$\Delta(uv) = v\Delta u + u\Delta v + (\Delta u)(\Delta v).$$

当量 Δu 和 Δv 非常小时, 那个小区域的面积事实上会非常非常小, 基本上可以忽略不计. 因此,

$$\Delta(uv) \approx v\Delta u + u\Delta v.$$

如果将上式除以 Δx, 然后取极限, 近似符号就会变成直等号, 会得到乘积法则

$$\frac{\mathrm{d}}{\mathrm{d}x}(uv) = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x}.$$

事实上, 这非常接近于真正的证明!

开始讲解链式求导法则之前, 先来证明一下三个函数的乘积法则. 正如我们之前看到的, 它由下式给出:

$$\frac{\mathrm{d}}{\mathrm{d}x}(uvw) = \frac{\mathrm{d}u}{\mathrm{d}x}vw + u\frac{\mathrm{d}v}{\mathrm{d}x}w + uv\frac{\mathrm{d}w}{\mathrm{d}x}.$$

这里的小窍门是令 $z = vw$, 这样 uvw 就是 uz. 首先, 对于 $z = vw$ 可以使用乘积法则:

$$\frac{\mathrm{d}z}{\mathrm{d}x} = w\frac{\mathrm{d}v}{\mathrm{d}x} + v\frac{\mathrm{d}w}{\mathrm{d}x}.$$

然后, 对 uz 使用乘积法则, 得到

$$\frac{\mathrm{d}}{\mathrm{d}x}(uvw) = \frac{\mathrm{d}}{\mathrm{d}x}(uz) = z\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}z}{\mathrm{d}x}.$$

剩下要做的就是用 vw 替换 z 以及用上式替换 $\mathrm{d}z/\mathrm{d}x$, 得到

$$\frac{\mathrm{d}}{\mathrm{d}x}(uvw) = z\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}z}{\mathrm{d}x} = vw\frac{\mathrm{d}u}{\mathrm{d}x} + u\left(w\frac{\mathrm{d}v}{\mathrm{d}x} + v\frac{\mathrm{d}w}{\mathrm{d}x}\right).$$

将上式展开, 就可以得到想要的公式了.

最后, 来考虑一下链式求导法则. 假设 $y = f(u)$ 及 $u = g(x)$. 这意味着, u 是 x 的函数, y 是 u 的函数. 如果将 x 稍作改变, 结果是 u 也会有相应的变化. 因为如此, y 也会改变. 那 y 将有多大的改变呢?

好吧, 让我们从关注函数 u 开始, 观察对于 x 的一个小的变化它是如何反应的. 忆及 $u = g(x)$, 因此正如 5.2.7 节所讨论的, u 的变化可以近似看成 $g'(x)$ 乘以 x 的变化. 你可以将 $g'(x)$ 看作是一种拉伸因子. (例如, 如果你站在游乐园中那些可以让你变高变瘦两倍的哈哈镜前面, 然后踮起脚尖, 你的镜像将变高两倍于你实际身高的高度.) 这可描述为

$$\Delta u \approx g'(x)\Delta x.$$

现在可以对用 u 表达的 y 来重复以上练习. 由于 $y = f(u)$, u 的一个变化会引起 y 中的近似 $f'(u)$ 倍于此的一个变化:

$$\Delta y \approx f'(u)\Delta u.$$

将这两个式子写在一起, 得到

$$\Delta y \approx f'(u)g'(x)\Delta x.$$

因此, x 的变化首先被因子 $g'(x)$ 拉伸了, 然后又被因子 $f'(u)$ 拉伸了. 总体的效果就是被两个拉伸因子 $f'(u)$ 和 $g'(x)$ 的乘积拉伸了. (毕竟, 如果你将一片口香糖拉伸两倍, 然后将**被拉伸过的口香糖**再拉伸三倍, 这与将原始的那片口香糖拉伸六倍是一样的.) 最后一个式子暗示了

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(u)g'(x).$$

从这里, 你不用太费劲就可以得到链式求导法则的两种形式中的任意一种. 为了得到第一种形式, 忆及 $u = g(x)$ 及 $y = f(u)$, 得到 $y = f(g(x))$; 然后, 令 $y = h(x)$ 并将上式重写为

$$h'(x) = f'(u)g'(x) = f'(g(x))g'(x).$$

为了得到第二种形式, 我们将 $f'(u)$ 解释为 $\mathrm{d}y/\mathrm{d}u$, 将 $g'(x)$ 解释为 $\mathrm{d}u/\mathrm{d}x$, 于是以上关于 $\mathrm{d}y/\mathrm{d}x$ 的式子就转化为

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}.$$

虽然上述解释不是正式的证明, 但它已经相当接近了.

6.3　求切线方程

那么, 求导又有什么用处呢? 一个好处就是, 可以使用导数来求所给曲线的切线方程. 假设有一条曲线 $y = f(x)$ 和曲线上一个特定的点 $(x, f(x))$, 那么过该点的切线的斜率是 $f'(x)$, 并且此切线通过点 $(x, f(x))$. 现在, 就可以使用点斜式来求切线方程了. 具体细节如下:

(1) **求斜率**, 通过求导函数并代入给定的 x 值;

(2) **求直线上的一点**, 通过将给定的 x 值代入原始函数本身得到 y 坐标, 将坐标写在一起并称之为点 (x_0, y_0); 最后,

(3) **使用点斜式** $y - y_0 = m(x - x_0)$ 来求方程.



这里有个例子. 令 $y = \left(x^3 - 7\right)^{50}$. 该函数图像在 $x = 2$ 处的切线方程是什么呢? 首先需要导函数. 需要使用链式求导法则. 令 $u = x^3 - 7$, 因此 $y = u^{50}$. 然后有 $\mathrm{d}y/\mathrm{d}u = 50u^{49}$ 及 $\mathrm{d}u/\mathrm{d}x = 3x^2$. 根据链式求导法则,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = 50u^{49} \times 3x^2 = 150x^2(x^3 - 7)^{49}.$$

(记住, 需要用 $x^3 - 7$ 替换 u 以便让一切都用 x 表示.) 现在需要代入 $x = 2$. 对于 x 的这个值, 有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 150(2)^2(2^3 - 7)^{49} = 150 \times 4 \times 1^{49} = 600.$$

很好, 我们找到了要找的切线的斜率. 接下去, 需要求它通过的那一点, 也就是把

$x=2$ 代入原始函数并看看 y 是什么. 事实上, $y=\left(2^3-7\right)^{50}=1^{50}=1$. 因此, 切线通过点 $(2,1)$. 使用点斜式, 可以看到切线方程是 $(y-1)=600\,(x-2)$, 或者重写为 $y=600x-1199$. 这就是求切线所有步骤!

6.4 速度和加速度

求导的另一个应用是计算运动物体的速度和加速度. 在 5.2.2 节中, 我们想象了一个物体沿着实轴运动. 我们发现, 如果在时刻 t 它的位置是 x, 那么它在时刻 t 的速度[1]就是

$$\boxed{\text{速度}=v=\frac{\mathrm{d}x}{\mathrm{d}t}.}$$

现在, 正如速度是位置的瞬时变化比率, 物体的**加速度**是速度的瞬时变化比率. 也就是说, 加速度是速度关于时间 t 的导数. 由于速度是位置的导数, 我们发现, 加速度实际上是位置的二阶导. 因此, 有

$$\boxed{\text{加速度}=a=\frac{\mathrm{d}v}{\mathrm{d}t}=\frac{\mathrm{d}^2x}{\mathrm{d}t^2}.}$$



例如, 假设一个物体在时刻 t 的位置由 $x=3t^3-6t^2+4t-2$ 给出, 其中 x 的单位是英尺, t 的单位是秒. 在时刻 $t=3$ 时该物体的速度和加速度分别是什么呢? 通过对位置关于时间求导得到速度：$v=\mathrm{d}x/\mathrm{d}t=9t^2-12t+4$. 现在再对这个新的表达式关于时间求导得到加速度：$a=\mathrm{d}v/\mathrm{d}t=18t-12$. 代入 $t=3$, 得到 $v=9\,(3)^2-12\,(3)+4=49$ 英尺/秒, 及 $a=18\,(3)-12=42$ 英尺/秒^2.

为什么加速度的单位是英尺每二次方秒呢? 其实, 当问一个物体的加速度是什么的时候, 实际上是在问该物体的速率变化有多快. 如果在一个为期 2 秒的时间段里, 速率由 15 英尺/秒变成 25 英尺/秒, 那么它的 (平均) 变化是 5 英尺/秒. 因此, 加速度的单位应该是英尺每二次方秒. 一般来说, 处理加速度的时候, 总是需要把时间单位平方.

负常数加速度

假设将一个球径直上抛, 它会上升然后落回 (除非它撞到了某物或被某人抓住了). 这是因为地球的引力将其拉向地面. 牛顿 (微积分的先驱之一) 认识到该力的效果就是：该球带有常数加速度向下运动. (假设没有空气阻力.)

由于该球先上升然后下降, 最好调整一下数轴的方向, 使它指向上下. 我们设 0 点为地面, 并且向上为正. 由于加速度是向下的, 它一定是一个负的量, 同时由于它是常数, 可以称之为 $-g$. 在地球上, g 大约是 9.8 米每二次方秒, 但在月球上这

[1] 从现在开始, 我们不再加 "瞬时" 两字, "速度" 将总是指瞬时速度, 除非明确说是 "平均速度".

个量会小得多. 不管怎样, 如果要理解这个球是如何运动的, 需要知道在时刻 t 它的位置和速度.

让我们从速度开始. 我们知道 $a = \mathrm{d}v/\mathrm{d}t$. 在上一节的例子中, 我们知道 v 是什么, 于是对其求导得到了 a. 不幸的是, 这一次恰恰反了过来, 我们知道 a(它就是常数 $-g$) 且需要求出 v. 一旦我们知道了 v, 同样的情况也会发生在 x 上. 在这两种情况下, 都需要反转微分的过程. 不幸的是, 我们对此还没有准备好 (那是积分部分所要讲的内容). 因此, 现在我只想告诉你答案, 然后通过微分来验证它：

> 在时刻 $t = 0$ 从初始高度 h 被抛出的带有初始速度 u 的一个物体, 满足以下方程：
> $$a = -g, \quad v = -gt + u \quad \text{及} \quad x = -\frac{1}{2}gt^2 + ut + h.$$

检验这些方程的相容性并不难. 关于 t 求导, 会看到 $\mathrm{d}v/\mathrm{d}t = -g$, 它就等于 a; 以及 $\mathrm{d}x/\mathrm{d}t = -gt + u$, 它就是 v. 因此, $a = \mathrm{d}v/\mathrm{d}t$ 且 $v = \mathrm{d}x/\mathrm{d}t$. 同时, 当 $t = 0$ 时, $v = u$ 及 $x = h$. 这意味着, 初始速度是 u, 初始高度是 h. 一切都得到了验证.

现在, 来看一个如何使用上述公式的例子吧. 假设你从距离地面高度为 2 米的地方以 3 米/秒的速率向上抛一个球. 取 g 为 10 米每二次方秒, 我们想要知道五点：

(1) 需要多久该球撞到地面?

(2) 当该球撞击地面时, 其运动有多快?

(3) 该球能上升到多高?

(4) 如果以相同的速率向下抛球, 需要多久该球撞到地面?

(5) 在那种情况下, 当它撞击地面时, 其运动有多快?

在原始情形下, 我们知道 $g = 10$, 初始高度为 $h = 2$, 初始速度为 $u = 3$. 这意味着以上公式变成

$$a = -10, \quad v = -10t + 3, \quad x = -\frac{1}{2}(10)t^2 + 3t + 2 = -5t^2 + 3t + 2.$$

对于第一点, 要求出需要多久该球撞击地面. 这显然只有其高度为 0 时才会发生. 因此, 设 $x = 0$, 求 t; 得到 $0 = -5t^2 + 3t + 2$. 如果将它因式分解为 $-(5t + 2)(t - 1)$, 就可以发现方程的解是 $t = 1$ 或 $t = -2/5$. 很明显第二个答案是不切合实际的, 在你还没有抛出之前该球不可能撞击地面! 因此, 答案一定是 $t = 1$. 也就是说, 抛出 1 秒后该球撞击地面.

对于第二点, 需要求出该球撞击地面时的速率. 没问题, 我们知道 $v = -10t + 3$, 并且现在知道当 $t = 1$ 时该球撞击地面. 将其代入, 得到 $v = -10 + 3 = -7$. 因此, 该球撞击地面时的速度是 -7 米/秒. 为什么是负的? 因为该球撞击地面时它是向下运动的, 向下的为负. 该球的速率就是速度的绝对值, 或 7 米/秒.

为了求解第三点, 你需要意识到, 当速度为 0 时该球达到它路径的最高点. 在向上的过程中, 速度是正的; 在向下的过程中, 速度是负的; 而当该球从向上变为向下运动时, 其速度一定是 0. 那么何时 v 等于 0 呢? 我们只需要求解 $-10t+3=0$. 答案是 $t=3/10$. 也就是说, 在抛出该球之后的 0.3 秒它达到其路径的最高点. 那么有多高呢? 只需要将 $t=3/10$ 代入公式 $x=-5t^2+3t+2$ 就可以得到

$$x=-5\left(\frac{3}{10}\right)^2+3\left(\frac{3}{10}\right)+2=\frac{49}{20}.$$

也就是说, 这时该球下距地面 49/20 米.

对于最后两点, 你是将球向下抛出. 我们仍然有 $g=10$ 及初始高度 $h=2$, 但初始速度 u 是什么呢? 不要错认为 u 仍然为 3! 由于球向下抛出, 初始速度是负的. 向下 3 米/秒的速率对应于初始速度 $u=-3$. 忽略这个负号是个常见的错误, 因此一定要警惕. 不管怎样, 方程现在变成

$$a=-10,\quad v=-10t-3\quad \text{及}\quad x=-\frac{1}{2}(10)t^2-3t+2=-5t^2-3t+2.$$

注意到, 这些方程和将球向上抛出情景下的方程很相似. 为了求解该问题的第四点, 需要求出该球撞击地面的时刻. 正如在第一点中所做的, 设 $x=0$, 然后有 $0=-5t^2+3t+2=-(5t-2)(t+1)$. 因此, $t=2/5$ 或 $t=-1$. 这一次舍弃 $t=-1$, 因为它是在抛球之前. 因此, 一定有 $t=2/5$. 也就是说, 在抛出后的 0.4 秒该球撞击地面. 它小于向上抛球时该球撞击地面所用的时间 (那是 1 秒), 这是自然的, 因为该球不需要先上升然后再下降. 对于最后一点, 要知道该球撞击地面时运动有多快; 因此, 将 $t=2/5$ 代入速度的公式, 得到 $v=-10(2/5)-3=-4-3=-7$. 再一次地, 该球以 7 米每秒的速率撞击地面. 有趣的是, 不管是将球向上抛还是向下抛 (只要它是以相同的速率从同一高度抛出), 它都以相同的速率撞击地面, 只是所用的时间有所不同.

6.5 导数伪装的极限



运动暂时讲得够多了. 现在考虑如何来求解极限

$$\lim_{h\to 0}\frac{\sqrt[5]{32+h}-2}{h}.$$

这看起来一点希望都没有. 甚至同乘以共轭表达式 $\sqrt[5]{32+h}+2$ 的技巧也不起作用, 因为它是五次方根, 不是平方根. (你自已试着看一下!) 因此, 暂且将它放在一边并考虑一个相关的极限

$$\lim_{h\to 0}\frac{\sqrt[5]{x+h}-\sqrt[5]{x}}{h}.$$

注意到这里的虚拟变量是 h 而不是 x. 这个极限看起来也很难解决, 但或许它让你感觉有点似曾相识. 它和以下公式中的极限非常相似：

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x).$$

因此, 所要做的就是设 $f(x)=\sqrt[5]{x}$, 并且注意到 $f'(x)=\frac{1}{5}x^{-4/5}$. (为了求导, 我们将 $\sqrt[5]{x}$ 写作 $x^{1/5}$.) 导数方程变为

$$\lim_{h\to 0}\frac{\sqrt[5]{x+h}-\sqrt[5]{x}}{h}=\frac{1}{5}x^{-4/5}.$$

因此, 等号左边的极限其实是一个伪装的导数! 我们需要创造一个函数 f 并对它求导来求此极限.

现在, 可以返回到初始极限

$$\lim_{h\to 0}\frac{\sqrt[5]{32+h}-2}{h}.$$

这其实是我们刚刚求解的极限

$$\lim_{h\to 0}\frac{\sqrt[5]{x+h}-\sqrt[5]{x}}{h}=\frac{1}{5}x^{-4/5}$$

的一个特例. 如果在此极限中设 $x=32$, 就会得到

$$\lim_{h\to 0}\frac{\sqrt[5]{32+h}-\sqrt[5]{32}}{h}=\frac{1}{5}\times 32^{-4/5}.$$

由于 $\sqrt[5]{32}=2$ 及 $32^{-4/5}=1/16$, 这就证明了

$$\lim_{h\to 0}\frac{\sqrt[5]{32+h}-2}{h}=\frac{1}{5}\times 32^{-4/5}=\frac{1}{5}\times\frac{1}{16}=\frac{1}{80}.$$

不要误以为这很简单. 这是一个双重的伪装：不仅是在处理一个导数, 实际上还是在计算一个特定点 (在这里是 32) 上的导数. 你最好先解得一般情况, 然后再代入 x 具体的值. 这里是另一个例子：

$$\lim_{h\to 0}\frac{\sqrt{(4+h)^3-7(4+h)}-6}{h}.$$

可以通过用共轭表达式和分子分母相乘来求解, 但它也是一个伪装的导数. 由于处理的是 $4+h$, 因而试着用 x 替换 4. 分子中的第一项变为 $\sqrt{(x+h)^3-7(x+h)}$. 这暗示着或许可以试着设 $f(x)=\sqrt{x^3-7x}$. 在 6.2.5 节中, 我们已经看到 $f'(x)=(3x^2-7)/2\sqrt{x^3-7x}$, 因此方程

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x)$$

变为

$$\lim_{h\to 0}\frac{\sqrt{(x+h)^3-7(x+h)}-\sqrt{x^3-7x}}{h}=\frac{3x^2-7}{2\sqrt{x^3-7x}}.$$

最后, 如果将 $x=4$ 代入并化简 (注意到 $\sqrt{x^3-7x}=\sqrt{64-28}=\sqrt{36}=6$), 会得到

$$\lim_{h\to 0}\frac{\sqrt{(4+h)^3-7(4+h)}-6}{h}=\frac{3(4)^2-7}{2(6)}=\frac{41}{12}.$$

如果求解一个极限有困难, 那它或许是一个伪装的导数. 迹象就是, 虚拟变量本身在分母上, 并且分子是两个量的差. 即使不是这样的, 仍然有可能是在处理一个伪装的导数. 例如,

$$\lim_{h\to 0}\frac{h}{(x+h)^6-x^6}$$

在分子上有一个虚拟变量. 这不要紧, 只需把它颠倒过来并先求出极限

$$\lim_{h\to 0}\frac{(x+h)^6-x^6}{h}.$$

为了求解, 设 $f(x)=x^6$, 则 $f'(x)=6x^5$. 因而有

$$\lim_{h\to 0}\frac{(x+h)^6-x^6}{h}=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x)=6x^5.$$

现在把它再颠倒一次, 得到

$$\lim_{h\to 0}\frac{h}{(x+h)^6-x^6}=\frac{1}{6x^5}.$$

我们将来 (确切地说, 是第 9 章和第 17 章) 会看到其他一些导数伪装的极限的例子. 所以要睁大眼睛：许多极限都是伪装的导数, 而你的工作就是揭开它们的伪装.[1]

6.6 分段函数的导数



考虑以下分段函数 f：

$$f(x)=\begin{cases}1 & \text{如果 } x\leqslant 0,\\ x^2+1 & \text{如果 } x>0.\end{cases}$$

① 事实上, 如果使用洛毕达法则 (参见第 14 章), 你经常甚至不需要去识别一个极限是否是一个伪装的导数.

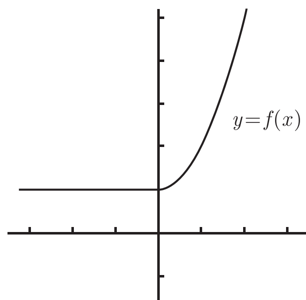


图　6-3

这个函数可导吗？让我们画出其图像来看看 (图 6-3). 这看起来相当平滑 —— 没有尖角. 事实上, 很明显, 除了可能在 $x=0$ 点上不可导, 函数 f 处处可导. 在 $x=0$ 的左侧, 函数 f 继承了常数函数 1 的可导性; 在 $x=0$ 的右侧, 函数 f 继承了 x^2+1 的可导性. 问题是, 在 $x=0$ 的两段接口处上发生了什么？

首先要检验函数在那里确实是连续的. 正如我们在 5.2.11 节看到的, 没有连续性就不可能有可导性. 为了确认 f 在 $x=0$ 上连续, 我们需要证明 $\lim\limits_{x\to 0} f(x)=f(0)$. 首先从 f 的定义, 可以看到 $f(0)=1$. 至于极限, 让我们将它分成左极限和右极限. 对于左极限, 当 x 在 0 的左侧时, 由于 $f(x)=1$, 有

$$\lim_{x\to 0^-} f(x)=\lim_{x\to 0^-}(1)=1.$$

至于右极限, 当 x 在 0 的右侧时, 由于 $f(x)=x^2+1$,

$$\lim_{x\to 0^+} f(x)=\lim_{x\to 0^+}(x^2+1)=0^2+1=1.$$

因此, 左极限等于右极限, 这意味着双侧极限存在并且等于 1. 这和 $f(0)$ 相等, 因此证明了 f 在 $x=0$ 上连续. (注意到, 对于左极限和右极限, 实际上只需将 $x=0$ 代入到适当的 f 的段中来求极限.)

我们仍然需要证明 f 在 $x=0$ 上可导. 为了求证, 必须证明在 $x=0$ 上的左导数和右导数相等 (回顾 5.2.10 节来回忆一下左导数和右导数的概念). 在 0 的左侧, 有 $f(x)=1$, 因此这时 $f'(x)=0$. 事实表明, 可以往右至 $x=0$, 得出

$$\lim_{x\to 0^-} f'(x)=\lim_{x\to 0^-} 0=0.$$

这表明 f 在 $x=0$ 上的左导数是 0. (更多详情参见附录 A 中的 A.6.10 节.) 在 0 的右侧, 有 $f(x)=x^2+1$, 因此 $f'(x)=2x$. 再一次地, 可以往左至 $x=0$:

$$\lim_{x\to 0^+} f'(x)=\lim_{x\to 0^+} 2x=2\times 0=0.$$

因此, f 在 $x=0$ 上的右导数是 $2\times 0=0$. 由于在 $x=0$ 上的左导数和右导数相等, 函数在 $x=0$ 上可导.

因此, 检验一个分段函数在分段连接点上是否可导, 需要检验分段在连接点上极限相等 (以证明连续性) **以及**分段的导数在连接点上相等. 否则, 在连接点上不可导.[①] 如果有两个以上的分段, 就必须在所有的连接点上检验连续性和可导性.

再来看一个有关求分段函数的导数的例子. 假设

① 事实上, 可导还要求导数在连接点上的左右极限都存在且有限. 有关的例子请参见 7.2.3 节.

$$g(x)=\begin{cases}|x^2-4| & \text{如果 } x\leqslant 1,\\ -2x+5 & \text{如果 } x>1.\end{cases}$$

g 在哪里可导呢？你或许会认为唯一的问题是在连接点 $x=1$ 上, 但事实上绝对值让事情变得更复杂了. 回想一下, 绝对值函数实际上是一个伪装的分段函数! 特别地, 当 $x\geqslant 0$ 时, $|x|=x$, 但当 $x<0$ 时, $|x|=-x$. 因此, 有

$$|x^2-4|=\begin{cases}x^2-4 & \text{如果 } x^2-4\geqslant 0,\\ -(x^2-4) & \text{如果 } x^2-4<0.\end{cases}$$

事实上, 不等式 $x^2-4<0$ 可以被重写为 $x^2<4$, 这意味着 $-2<x<2$. (除了更显然的 $x<2$, 注意别落了 $-2<x$!) 因此, 我们稍微化简得到

$$|x^2-4|=\begin{cases}x^2-4 & \text{如果 } x\geqslant 2 \text{ 或 } x\leqslant -2,\\ -x^2+4 & \text{如果 } -2<x<2.\end{cases}$$

由于在上述 $g(x)$ 的定义中, 项 $|x^2-4|$ 只有当 $x\leqslant 1$ 才出现, 因此可以将一切拼起来并去掉绝对值, 重新将 $g(x)$ 写成

$$g(x)=\begin{cases}x^2-4 & \text{如果 } x\leqslant -2,\\ -x^2+4 & \text{如果 } -2<x\leqslant 1,\\ -2x+5 & \text{如果 } x>1.\end{cases}$$

因此, 事实上有两个连接点：$x=-2$ 和 $x=1$. 由于组成 g 的三个分段都是处处可导, 我们知道, 除了可能在连接点上不可导外, g 本身处处可导. 让我们检验一下连接点上的性质, 我们由 $x=-2$ 开始. 首先是连续性. 从左侧, 有

$$\lim_{x\to(-2)^-}g(x)=\lim_{x\to(-2)^-}x^2-4=(-2)^2-4=0;$$

而从右侧, 有

$$\lim_{x\to(-2)^+}g(x)=\lim_{x\to(-2)^+}-x^2+4=-(-2)^2+4=0.$$

由于两个极限相等, 因此 g 在 $x=-2$ 上连续. 现在检验导数. 对于左导数, 有

$$\lim_{x\to(-2)^-}g'(x)=\lim_{x\to(-2)^-}2x=2(-2)=-4;$$

而对于右导数, 有

$$\lim_{x\to(-2)^+}g'(x)=\lim_{x\to(-2)^+}-2x=2(-2)=4.$$

由于它们不相等, 故函数 g 在 $x=-2$ 上不可导.

在另外一个连接点 $x=1$ 上又如何呢？重复之前的步骤. 左连续：

$$\lim_{x\to1^-} g(x) = \lim_{x\to1^-} -x^2+4 = -(1)^2+4 = 3;$$

右连续：

$$\lim_{x\to1^+} g(x) = \lim_{x\to1^+} -2x+5 = -2(1)+5 = 3.$$

它们相等, 因此 g 在 $x=1$ 上连续. 现在, 左可导性：

$$\lim_{x\to1^-} g'(x) = \lim_{x\to1^-} -2x = -2(1) = -2;$$

右可导性：

$$\lim_{x\to1^+} g'(x) = \lim_{x\to1^+} -2 = -2.$$

由于它们相等, 因此函数 g 在 $x=1$ 上可导.

我们已经回答了原始问题, 但不管怎样, 还是要画出图像来看看到底发生了什么. 为了画出 $y=|x^2-4|$ 的图像, 要先画 $y=x^2-4$ 的图像. 这是一个抛物线, 其 x 轴截距在 2 和 -2 (那里就是 $y=0$ 的地方) 并且其 y 轴截距在 -4. 为了得到绝对值, 将 x 轴下方的一切关于 x 轴做反射. 翻转的那部分是曲线 $y=-x^2+4$ 的一部分. 最后, 直线 $y=-2x+5$ 有 y 轴截距 5 及 x 轴截距 5/2, 因此并不难画出图像. 在图 6-4 的两幅图中, 左图显示了组成 $g(x)$ 的所有的函数, 右图则取我们所需, 也就是 $y=g(x)$ 的图像.

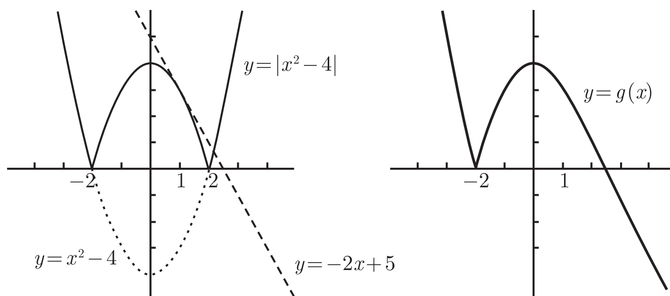


图　6-4

它实际上也是看起来处处连续且处处可导, 除了在尖角 $(-2,0)$ 外. 特别地, 在连接点 $x=1$ 上一切正常, 正如我们计算的那样.

6.7　直接画出导函数的图像

假设有一个函数的图像, 你不知道它的方程, 但又想要画出其导函数的图像. 这时公式和法则帮不上你, 你需要的是对微分有一个很好的理解.

下面是基本思想. 将函数的图像想象成一座山, 并想象有一个小登山者在从左到右地爬上爬下. 在攀登的每一点上, 登山者会大声地喊出他认为攀登有多困难.

如果地形平坦, 登山者会大声喊出数 0 以表示难度. 如果地形呈现向上的斜坡, 登山者会大声喊出一个正的数; 攀登越陡峭, 数越大. 如果地形呈现向下的斜坡, 那么攀登实际上很轻松, 因此难度是负的. 也就是说, 登山者会大声喊出一个负的数. 向下的斜坡越陡越轻松, 因此数会越来越负. (如果下坡确实非常陡, 那它或许会让下行变得更不安全, 但它显然也让下降变得更为快速!)

这里的要点是：山的高度本身不重要, 重要的是陡峭程度. 特别地, 你可以将整个图像向上平移, 登山者还是会大声喊出相同的难度程度来. 这意味着, 如果你从一个函数的图像画一个导函数的图像, 该函数的 y 轴截距是不重要的!

来看一个例子：画出下述让人恐惧的函数的导函数的图像, 如图 6-5 所示.

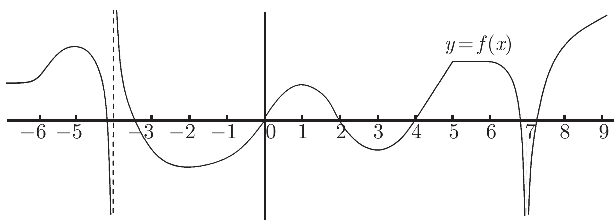


图 6-5

不要惊慌. 只需在各个不同的点上画一个小登山者并想象登山者在每一点上大声喊出难度程度. 然后, 你所要做的就是在另一套坐标上画出这些难度程度. 特别要留心的是那些路径平坦的点; 这可以出现在一个长的平坦的区域中 (如上图中的 $x = 5$ 和 $x = 6$ 之间), 或者在一个峰的顶部 (如在 $x = -5$ 或 $x = 1$), 又或者在一个谷的底部 (如在 $x = -2$ 或 $x = 3$). 那里你肯定是要画出登山者的. 图 6-6 是在一些位置放上登山者的 f 的图像.

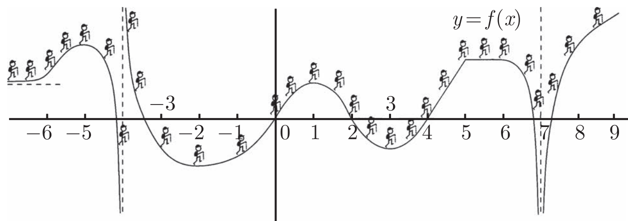


图 6-6

现在, 为导函数的图像来画一套坐标. y 轴标记为 “难度程度”, 从上至下由难至易. 然后, 基于小登山者大声喊出的难度程度, 你应该能够用铅笔描出一些点来. 回想一下, 登山者并不关心山有多高, 他**只关心山有多陡**! 基于此, 会得到图 6-7 上的一些点.

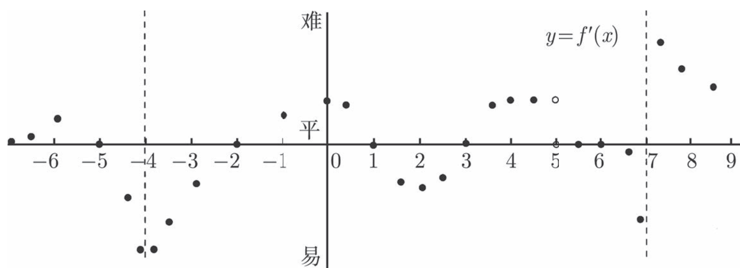


图　6-7

以下是对于如何得出这些点的详细解释.

- 在 $y = (x)$ 的图像的最左侧, 登山者开始只是缓缓地上坡. 因此, 将画出一些高度稍高于 0 的一些点.
- 往前走, 走到 $x = -6$, 登山者开始上坡, 因此难度上升, 这些点也变高了 (更难了).
- 然后, 开始变得略微容易点, 直到 $x = -5$ 时, 登山者达到峰的顶部, 那里是平坦的. 特别地, 当 $x = -5$ 时, 导函数有一个 x 轴截距.
- 在 $x = -5$ 之后, 原始函数的曲线开始变成下坡, 首先较为平缓, 然后越来越陡. 这意味着, 攀登将变得越来轻松, 直到它变得非常非常轻松. 因此, 导函数在 $x = -4$ 处有一条垂直渐近线.
- 在该渐近线的另外一侧, 攀登也很容易, 因为登山者将下坡, 开始非常陡, 并在 $x = -2$ 处到达谷底. 因此, 在导函数曲线上, 垂直渐近线实际上始于 $-\infty$ (非常非常容易) 并在 $x = -2$ 处爬升至 0. (在 $x = -5$ 和 $x = -4$ 之间有 x 轴截距以及在 $x = -4$ 和 $x = -3$ 之间也有, 但这都无关紧要. 原始函数的 x 轴截距不重要.)
- 在 $x = -2$ 到达谷底之后, 登山者必须上坡一会儿, 因此攀登变困难了. 尽管在 $x = 0$ 之后变得略微容易点, 但他依然要往上爬, 直到 $x = 1$ 的山顶. 这意味着, 导函数的曲线上升, 直到 $x = 0$, 然后下降, 直到 $x = 1$, 得到一个 x 轴截距.
- 在走向 $x = 3$ 处谷底的路上, 情况发生了逆转：下坡越来越陡, 直到 $x = 2$, 然后坡度减缓些, 但仍然是下坡. 因此, 导函数的曲线下降, 在 $x = 2$ 处达到一个最小值, 然后上升, 直到 $x = 3$, 得到一个 x 轴截距.
- 从 $x = 3$ 处的谷底起, 攀登一直都很困难, 直到 $x = 4$. 然而, 在 $x = 4$ 和 $x = 5$ 之间, 攀登的难度是均匀的, 因为斜率是常数. 因此, 导函数的曲线从 $x = 3$ 上升, 直到 $x = 4$, 然后在 $x = 4$ 和 $x = 5$ 之间, 保持在同一高度 (难度程度).
- 在 $x = 5$, 斜率突然地改变了. 在没有任何预警的情况下, 它突然变平坦了,

然后保持这种平坦直到 $x = 6$. 因此, 导函数的曲线必须下降至 0 并且保持为 0 直到 $x = 6$. 导函数在 $x = 5$ 处有一个不连续点.

- 在 $x = 6$ 之后, 登山者发现, 随着曲线逼近 $x = 7$ 处的垂直渐近线, 攀登越来越容易了. 导函数的曲线在那里也有一条垂直渐近线.
- 在这条垂直渐近线的右侧, 攀登极度困难, 但当 x 走向 9 时, 攀登变得略微容易点. 因此, 导函数的曲线在 $x = 7$ 的右侧始于非常高的地方, 然后当攀登越来越容易时, 它变得越来越低.

现在, 只需要把这些点连起来! 图 6-8 分别是 $y = f(x)$ 和 $y = f'(x)$ 的图像.

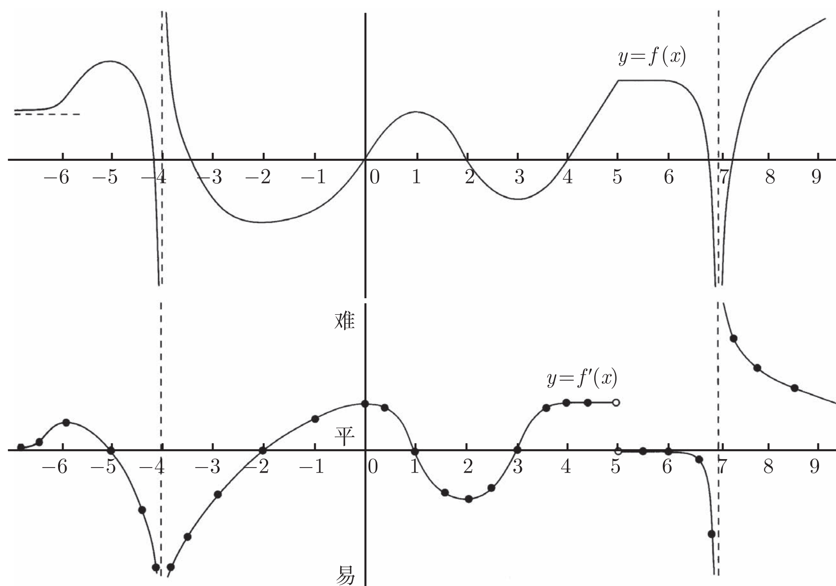


图 6-8

我们把所用的思想做一下总结.

- 当原始图像平坦时, 导函数的图像有一个 x 轴截距. 在上例中, 它们出现在 $x = -5$, $x = -2$, $x = 1$, $x = 3$ 及区间 $[5, 6]$ 的每一点上.
- 当原始图像的一部分是一条直线时, 导函数的图像是常数 (上例中, 它出现在区间 $[4, 5]$ 上).
- 如果原始图像有一条水平渐近线, 其导函数图像经常也有一条水平渐近线, 但如果是那样的话, 它将在 $y = 0$ 的高度, 而不是渐近线的原始高度上 (正如上例中图像的左端).
- 原始图像中的垂直渐近线通常导致导函数在相同位置上也有垂直渐近线[①],

① 如果一个函数有一条垂直渐近线, 那么它的导函数在相同的位置上也有一条垂直渐近线, 一般而言, 这实际上是不正确的. 一个例子是: $y = 1/x + \sin(1/x)$ 在 $x = 0$ 处. 你能看出为什么吗?

尽管方向可能会改变. 例如上例中, 在 $x = 7$ 处, 在渐近线的两侧, 原始函数的曲线都走向 $-\infty$, 但导函数却有相反的符号. 在 $x = -4$ 处的垂直渐近线也受到类似的影响.

如果有怀疑的话, 就请使用那些值得信赖的登山者进行验证吧!

第 7 章　三角函数的极限和导数

到目前为止, 我们讨论的大多数极限和导数问题只涉及多项式或多项式型函数. 现在让我们拓宽视野, 来看看三角函数的极限和导数吧. 特别地, 我们将关注以下几个方面：

- 三角函数在小数、大数以及其他变量值时的行为;
- 三角函数的导数;
- 简谐运动.

7.1　三角函数的极限

考虑两个极限

$$\lim_{x\to 0}\frac{\sin(5x)}{x} \quad \text{和} \quad \lim_{x\to\infty}\frac{\sin(5x)}{x}.$$

它们看上去几乎是一样的. 唯一的区别是, 第一个极限是在 $x \to 0$ 时取的, 而第二个则是在 $x \to \infty$ 时取的. 但正是这点区别, 造就了大不同! 正如我们即将看到的, 这两个极限的答案和求解技巧几乎没有共同点. 因此, 需要切实留意你是在非常小的数 (如上述第一个极限) 上还是在非常大的数 (如上述第二个极限) 上取正弦或余弦或正切的极限. 我们将分别考察这两种情况, 然后再看看当这两种情况都不适用时会发生什么.

在开始之前, 需要提醒的是, 你无法只通过看 $x \to 0$ 或 $x \to \infty$ 来了解自己处理的是哪种情况. 你需要知道自己是在哪里计算三角函数的值. 例如, 考虑两个极限

$$\lim_{x\to 0} x\sin\left(\frac{5}{x}\right) \quad \text{和} \quad \lim_{x\to\infty} x\sin\left(\frac{5}{x}\right).$$

在第一个极限中, 你要取的是 $5/x$ 的正弦, 当 x 接近于 0 时, 它实际上是一个巨大的数 (正的或负的, 取决于 x 的符号). 因此, 第一个极限根本就不是小数的情况, 它属于大数的情况! 类似地, 在第二个极限中, 当 x 非常大时, 量 $5/x$ 会非常小, 因此这才是真正的小数的情况. 在接下来的几节中, 我们会解答上述的所有四个极限.

7.1.1　小数的情况

我们知道 $\sin(0) = 0$. 那好, 当 x 接近于 0 时, $\sin(x)$ 看起来会怎样呢？当然, 如果那样的话, $\sin(x)$ 也会接近于 0, 但它距离 0 有多近呢？事实表明, $\sin(x)$ 与 x

本身近似相等!

例如, 如果用计算器, 将它设置为弧度模式, 并求 $\sin(0.1)$, 你得到的结果大约是 0.0998, 它非常接近于 0.1. 尝试一个更接近于 0 的数, 你就会发现, 选取的数的正弦值与选取的原始数值非常接近.

不妨看一下此情况的图像. 图 7-1 是 $y = \sin(x)$ 和 $y = x$ 在同一坐标系下的图像, 只取 x 在 -1 和 1 之间 (近似的) 的值.

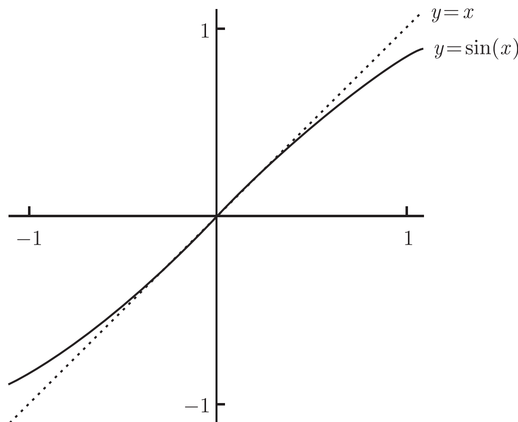


图　7-1

这两个图像非常相似, 尤其是当 x 接近于 0 的时候. (当然, 如果再多画一点 $y = \sin(x)$ 的图像, 我们就会看到熟悉的波形; 只有将它放大成这样的时候, 我们才会看到 $\sin(x)$ 多么接近于 x.) 因此, 我们可以说, 当 x 非常小的时候, $\sin(x)$ 接近于 x. 要是 $\sin(x)$ 实际上等于 x 的话, 那么

$$\frac{\sin(x)}{x} = 1$$

成立. 事实上, 上述等式永远都不会成立, 不过在 $x \to 0$ 时的极限中确实有：

$$\boxed{\lim_{x \to 0} \frac{\sin(x)}{x} = 1.}$$

这个公式非常重要. 基本上, 这是解决涉及三角函数的微积分问题的关键所在. 我们将在 7.2 节使用它来求三角函数的导数, 并会在 7.1.5 节对它进行证明.

$\cos(x)$ 又怎样呢? 好吧, $\cos(0) = 1$, 因此情况大为不同. 暂且说一个小数的余弦非常接近于 1. 我们写作

$$\boxed{\lim_{x \to 0} \cos(x) = 1.}$$

特别要注意的是, 不像之前那个涉及 $\sin(x)$ 的公式, 这里的分母中没有 x 的因子. 要将 x 的因子放在分母中会怎样呢? 我们很快就会看到, 不过我想先看看 $\tan(x)$.

这里的关键是, 将 $\tan(x)$ 写成 $\sin(x)/\cos(x)$. 分子是 $\sin(x)$, 当 x 非常小时, 它非常接近于 x. 另一方面, 这时分母接近于 1. 显而易见, 它们的比应该就好像 $x/1$, 即 x. 事实上也确实如此. 将 $\cos(x)$ 从分母中分离出来, 得到

$$\lim_{x\to 0}\frac{\tan(x)}{x}=\lim_{x\to 0}\frac{\frac{\sin(x)}{\cos(x)}}{x}=\lim_{x\to 0}\left(\frac{\sin(x)}{x}\right)\left(\frac{1}{\cos(x)}\right)=(1)\left(\frac{1}{1}\right)=1.$$

这样就证明了

$$\boxed{\lim_{x\to 0}\frac{\tan(x)}{x}=1.}$$

这意味着, 当 x 非常小时, $\sin(x)$ 和 $\tan(x)$ 的行为很相似, 但 $\cos(x)$ 却有所不同. 现在让我们来看看, 当 $x\to 0$ 时 $\cos(x)/x$ 会发生什么. 也就是说, 我们试图理解

$$\lim_{x\to 0}\frac{\cos(x)}{x}.$$



如果你只是将 $x=0$ 代入上式的话, 那么就会得到 $1/0$. 这意味着, $y=\cos(x)/x$ 的图像在 $x=0$ 处有一条垂直渐近线. 这看上去很像 x 很小时的 $1/x$; 特别地, 你应该意识到

$$\lim_{x\to 0^+}\frac{\cos(x)}{x}=\infty,\quad \lim_{x\to 0^-}\frac{\cos(x)}{x}=-\infty,\quad 所以\quad \lim_{x\to 0}\frac{\cos(x)}{x}\text{DNE}.$$

(回想一下, "DNE" 表示 "不存在".) 这确实与正弦或正切的情况有所不同.

7.1.2 问题的求解 —— 小数的情况



以下是一个简单的例子：求

$$\lim_{x\to 0}\frac{\sin(x^2)}{x^2}.$$

首先注意到当 x 接近于 0 时, x^2 也接近于 0, 因此实际上是在取一个小数的正弦. 现在, 我们知道极限

$$\lim_{x\to 0}\frac{\sin(x)}{x}=1$$

成立. 如果你用 x^2 (它是 x 的连续函数) 替换 x, 那么会得到下面的有效极限：

$$\lim_{x^2\to 0}\frac{\sin(x^2)}{x^2}=1.$$

这几乎是我们想要的极限了. 事实上, 只需注意到一点, 当 $x\to 0$ 时, $x^2\to 0$, 因此最后可以求得该极限为

$$\lim_{x\to 0}\frac{\sin(x^2)}{x^2}=1.$$

当然, x^2 没什么特别的; 当 $x=0$ 时, x 的任意其他连续函数都是 0. 特别地, 我们可以自然而然地知道极限

$$\lim_{x\to0}\frac{\sin(5x)}{5x}=1;\quad \lim_{x\to0}\frac{\sin(3x^7)}{3x^7}=1;\quad 甚至\quad \lim_{x\to0}\frac{\sin(\sin(x))}{\sin(x)}=1$$

成立. 用 "tan" 替换 "sin", 以上等式依然成立, 但千万别用 "cos"! 不管怎样, 我们可以总结如下：

$$\boxed{\lim_{x\to0}\frac{\sin(小数)}{同样的小数}=1}\quad 和 \quad \boxed{\lim_{x\to0}\frac{\tan(小数)}{同样的小数}=1.}$$

这里至关重要的是, 分母要与分子中正弦或正切的变量**相匹配**, 并且当 x 很小的时候, 这个量要很小. 当然, 对于余弦, 我们所能说的只有

$$\boxed{\lim_{x\to0}\cos(小数)=1.}$$

这里就不存在是否匹配的问题了!

现在, 让我们回头看一下本章开始时的例子：

$$\lim_{x\to0}\frac{\sin(5x)}{x}.$$

这里的问题是, 我们取的是 $5x$ 的正弦, 但在分母上只有 x. 这两个量不匹配. 不过不要紧, 用 $\sin(5x)$ 除以 $5x$, 这样就匹配了, 然后再乘以该量, 使得结果不变. 也就是说, 将 $\sin(5x)$ 重新写作

$$\frac{\sin(5x)}{5x}\times(5x).$$

这与我们在 4.3 节中求解有理函数的极限时所用的技巧几乎是一样的! 让我们来看看, 在这种情况下, 它是如何发挥作用的吧：

$$\lim_{x\to0}\frac{\sin(5x)}{x}=\lim_{x\to0}\frac{\dfrac{\sin(5x)}{5x}\times(5x)}{x}.$$

现在, 保留 $\sin(5x)/5x$ 不变, 但从其他两个因子中消去 x, 得到

$$\lim_{x\to0}\frac{\sin(5x)}{x}=\lim_{x\to0}\frac{\sin(5x)}{5x}\times5,$$

由于匹配了项 $5x$ (一个在分母上, 一个在正弦的变量中), 由前可知, 该分式的极限为 1, 因此总的极限是 5. 放到一起, 问题的解如下：

$$\lim_{x\to0}\frac{\sin(5x)}{x}=\lim_{x\to0}\frac{\dfrac{\sin(5x)}{5x}\times(5x)}{x}=\lim_{x\to0}\frac{\sin(5x)}{5x}\times5=1\times5=5.$$

现在来看一个更难的例子. 极限

$$\lim_{x\to0}\frac{\sin^3(2x)\cos(5x^{19})}{x\tan(5x^2)}$$

是什么？我们分别来看一下该表达式中的四个因子. 首先考虑 $\sin^3(2x)$. 这其实就是 $(\sin(2x))^3$ 的另外一种写法. 为了处理 $\sin(2x)$, 需要用 $2x$ 做除法和乘法; 而为了处理其立方, 需要用 $(2x)^3$ 做除法和乘法. 也就是说, 用

$$\frac{(\sin(2x))^3}{(2x)^3} \times (2x)^3$$

替换 $(\sin(2x))^3$. 那么 $\cos(5x^{19})$ 又如何呢？其实, 当 x 很小的时候, $5x^{19}$ 也很小, 因此, 就是在取一个小数的余弦. 极限的结果应该是 1, 因此不用对第二个因子进行操作.

在分母上, 我们有一个因子 x, 不能对它做任何操作. (我们也不想, 它实际上已经很容易处理了!) 还有一个因子 $\tan(5x^2)$. 我们对 $5x^2$ 做除法和乘法, 以便用

$$\frac{\tan(5x^2)}{5x^2} \times (5x^2)$$

替换 $\tan(5x^2)$. 将所有这些放在一起, 得到

$$\lim_{x\to 0} \frac{\sin^3(2x)\cos(5x^{19})}{x\tan(5x^2)} = \lim_{x\to 0} \frac{\left[\frac{(\sin(2x))^3}{(2x)^3} \times (2x)^3\right]\cos(5x^{19})}{x\left[\frac{\tan(5x^2)}{5x^2} \times (5x^2)\right]}.$$

现在, 将所有与三角函数的变量不匹配的 x 的幂次都提出来：分子中的项 $(2x)^3$ 和分母中的项 x 和 $5x^2$. 然后, 重新将 $(\sin(2x))^3/(2x)^3$ 写作 $(\sin(2x)/2x)^3$ 并化简, 可以看到极限变为

$$\lim_{x\to 0} \frac{\frac{(\sin(2x))^3}{(2x)^3} \cdot \cos(5x^{19})}{\frac{\tan(5x^2)}{5x^2}} \times \frac{(2x)^3}{x(5x^2)} = \lim_{x\to 0} \frac{\left(\frac{\sin(2x)}{(2x)}\right)^3 \cos(5x^{19})}{\frac{\tan(5x^2)}{5x^2}} \times \frac{8x^3}{5x^3}.$$

最后, 可以在分子分母中消去 x^3, 并取极限. 由于正弦和正切有相匹配的分子和分母, 还有 $\cos(\text{小数}) \to 1$, 因此极限就是

$$\frac{(1)^3(1)}{1} \times \frac{8}{5} = \frac{8}{5}.$$



下面是本章开头部分的另外一个例子：极限

$$\lim_{x\to\infty} x\sin\left(\frac{5}{x}\right)$$

是什么？我们可以看出, 这个例子的确属于本节内容, 因为当 x 很大时, 量 $5/x$ 会非常小. 因此, 我们使用相同的方法. 在这种情况下, 用 $\sin(5/x)$ 除以并乘以 $5/x$：

$$\lim_{x\to\infty} x\sin\left(\frac{5}{x}\right) = \lim_{x\to\infty} x \cdot \frac{\sin\left(\frac{5}{x}\right)}{\frac{5}{x}} \times \frac{5}{x}.$$

现在, 可以消去因子 x 并化简得到

$$\lim_{x\to\infty} 5 \times \frac{\sin(5/x)}{5/x}.$$

如果把 “小数” 用 $5/x$ 替换, 可以立即看到, 当 $x \to \infty$ 时, 这个大分式的极限是 1, 因此最终的结果就是 5.

我们也可能会面对涉及正割、余割或余切的三角函数的极限. 例如, 极限

$$\lim_{x\to 0} \sin(3x)\cot(5x)\sec(7x)$$

是什么？为了求解该极限, 最稳妥的猜测是, 换用余弦、正弦或正切来表示它：

$$\lim_{x\to 0}(\sin(3x))\left(\frac{1}{\tan(5x)}\right)\left(\frac{1}{\cos(7x)}\right).$$

现在, 可以对正弦和正切项使用标准的乘除法技巧, 同时忽略余弦项. 于是我们看到该极限等于

$$\lim_{x\to 0}\left(\frac{\sin(3x)}{3x}\times(3x)\right)\left(\frac{1}{\frac{\tan(5x)}{5x}\times(5x)}\right)\left(\frac{1}{\cos(7x)}\right).$$

现在, $(3x)$ 和 $(5x)$ 这两项可以消去公因子 x, 得到 $3/5$, 而所有其他分式的极限趋于 1, 因此整个极限就是 $3/5$.

有一点你必须非常小心：说当 x 非常小时, $\sin(x)$ 表现得就像 x, 这只有在乘积或商的语境中才成立. 例如, 极限

$$\lim_{x\to 0}\frac{x-\sin(x)}{x^3}$$

就不能用本章介绍的方法进行求解. 说 $\sin(x)$ 表现得像 x, 故 $x-\sin(x)$ 表现得像 0, 这是错误的. (事实上, 除了常数函数 0 本身, 没有函数表现得像 0!) 为了求解以上极限, 需要洛必达法则 (参见第 14 章) 或麦克劳林级数 (参见第 24 章). 另一方面, 下面这个类似难度的极限却是我们现在可以求解的：

$$\lim_{x\to 0}\frac{1-\cos^2(x)}{x^2}.$$

同样, 你不能说, 当 x 非常小时, $\cos(x)$ 表现得就像 1, 故 $1-\cos^2(x)$ 表现得就像 $1-1^2=0$. 因此, 我们使用 $\cos^2(x)+\sin^2(x)=1$ 来将分子重写为 $\sin^2(x)$：

$$\lim_{x\to 0}\frac{1-\cos^2(x)}{x^2}=\lim_{x\to 0}\frac{\sin^2(x)}{x^2}.$$

由于 $\sin^2(x)$ 是 $(\sin(x))^2$ 的另一种写法, 可以将极限重写为

$$\lim_{x\to 0}\frac{(\sin(x))^2}{x^2}=\lim_{x\to 0}\left(\frac{\sin(x)}{x}\right)^2.$$

该极限就是 $1^2=1$. 因此,

$$\lim_{x\to 0}\frac{1-\cos^2(x)}{x^2}=1.$$

换句话说, 当 x 非常小时, $1-\cos^2(x)$ 表现得就像 x^2, 而根本不像 0. 不管怎样, 让

我们使用同样的思想来求解其他一些极限：

$$\lim_{x\to 0}\frac{1-\cos(x)}{x^2} \quad 和 \quad \lim_{x\to 0}\frac{1-\cos(x)}{x}.$$

我们将使用同样聪明的技巧来求解这两个极限. 基本思想就是, 用 $1+\cos(x)$ 和分子分母分别相乘, 使分子变成 $1-\cos^2(x)$, 进而使它写成 $\sin^2(x)$. 在第一个例子中, 我们有

$$\begin{aligned}\lim_{x\to 0}\frac{1-\cos(x)}{x^2} &= \lim_{x\to 0}\frac{1-\cos(x)}{x^2}\times\frac{1+\cos(x)}{1+\cos(x)}.\\ &= \lim_{x\to 0}\frac{1-\cos^2(x)}{x^2}\times\frac{1}{1+\cos(x)} = \lim_{x\to 0}\frac{\sin^2(x)}{x^2}\times\frac{1}{1+\cos(x)}\\ &= \lim_{x\to 0}\left(\frac{\sin(x)}{x}\right)^2\times\frac{1}{1+\cos(x)} = 1^2\times\frac{1}{1+1} = \frac{1}{2}.\end{aligned}$$

这里使用的事实就是 $\cos(0)=1$. 第二个例子很相似：

$$\begin{aligned}\lim_{x\to 0}\frac{1-\cos(x)}{x} &= \lim_{x\to 0}\frac{1-\cos(x)}{x}\times\frac{1+\cos(x)}{1+\cos(x)}.\\ &= \lim_{x\to 0}\frac{1-\cos^2(x)}{x}\times\frac{1}{1+\cos(x)} = \lim_{x\to 0}\frac{\sin^2(x)}{x}\times\frac{1}{1+\cos(x)}.\end{aligned}$$

到这里, 可以用 x^2 和项 $\sin^2(x)$ 做除法和乘法, 但还有一个较为简便的求极限的方法：将 $\sin^2(x)$ 写成 $\sin(x)\times\sin(x)$, 并将其中一个 $\sin(x)$ 因子和分母中的因子 x 放到一起. 由于 $\sin(0)=0$, 极限变为

$$\lim_{x\to 0}\left(\sin(x)\times\frac{\sin(x)}{x}\times\frac{1}{1+\cos(x)}\right) = 0\times 1\times\frac{1}{1+1} = 0.$$

最后这个极限将在 7.2 节中很有用, 因此让我们总结一下并记住它：

$$\boxed{\lim_{x\to 0}\frac{1-\cos(x)}{x} = 0.}$$

关于小数的情况我们已经讨论得足够多了, 现在来看看如何处理三角函数在大数上的极限吧.

7.1.3 大数的情况

考虑极限

$$\lim_{x\to\infty}\frac{\sin(x)}{x}.$$

正如我们刚刚看到的, 如果 $x\to 0$ 而不是 ∞, 那么极限为 1. 这是因为当 x 非常小时, $\sin(x)$ 表现得就像 x. 但当 x 变得越来越大的时候, $\sin(x)$ 的行为又如何呢？它会在 -1 和 1 之间来回振荡. 因此, 当 x 变大时它没有表现得像谁. 于是我们常常只好求助于 $\sin(x)$ (以及 $\cos(x)$) 最简单的性质之一：

$$\text{对于任意的 } x, \boxed{-1 \leqslant \sin(x) \leqslant 1} \quad \text{和} \quad \boxed{-1 \leqslant \cos(x) \leqslant 1}.$$

这时应用三明治定理 (参见 3.6 节) 就相当方便了. 事实上, 我们在 3.6 节已经看到

$$\lim_{x \to \infty} \frac{\sin(x)}{x} = 0.$$

不妨马上回去重温一下证明.

回过头来, 还记得当 x 变小时, $\cos(x)$ 是有所不同的吗? 不像 $\sin(x)$ 和 $\tan(x)$, 它不像 x. 另一方面, 当 x 变大时, $\tan(x)$ 也有所不同. 对于 $\tan(x)$ 来说, 没有类似于上面方框中关于 $\sin(x)$ 和 $\cos(x)$ 的不等式. 这是因为当 x 变大时, $\tan(x)$ 有垂直渐近线并且永远不会停下来 (参见 2.3 节 $\tan(x)$ 的图像).

这儿有一个使用三明治定理的更难的例子：求

$$\lim_{x \to \infty} \frac{x\sin(11x^7) - \frac{1}{2}}{2x^4}.$$

直觉告诉我们, $\sin(11x^7)$ 这一项无足轻重, 因此分子其实与 x 相当. 而分母中的 x^4 应该较分子的 x 占压倒性优势, 因此当 $x \to \infty$ 时, 整个表达式应该是趋于 0 的. 为了证明这一点, 首先来看看分子. 我们知道任何数的正弦都在 -1 和 1 之间, 因此,

$$-1 \leqslant \sin(11x^7) \leqslant 1$$

成立. 不过分子不只是 $\sin(11x^7)$, 需要用 x 和它相乘然后再减去 1/2. 事实上, 对于任意的 $x > 0$, 可以对以上不等式的所有三 “方” 都作相同的操作, 得到

$$-x - \frac{1}{2} \leqslant x\sin(11x^7) - \frac{1}{2} \leqslant x - \frac{1}{2}.$$

(如果 $x < 0$, 也就是说, 当 $x \to -\infty$ 时, 那么用负的 x 做乘法意味着, 你必须将所有的小于或等于号反转变成大于或等于号. 不然的话, 会得到同一个解.) 不管怎样, 这解决了分子的问题. 但我们还需要除以分母. 由于 $2x^4 > 0$, 可以将以上不等式除以 $2x^4$, 得到

$$\frac{-x - \frac{1}{2}}{2x^4} \leqslant \frac{x\sin(11x^7) - \frac{1}{2}}{2x^4} \leqslant \frac{x - \frac{1}{2}}{2x^4}.$$

万事俱备, 只欠东风. 现在我把它留给你, 请使用 4.3 节中的方法证明, 当 $x \to \infty$ 时, 外层两项的极限均为 0, 即

$$\lim_{x \to \infty} \frac{-x - \frac{1}{2}}{2x^4} = 0 \quad \text{和} \quad \lim_{x \to \infty} \frac{x - \frac{1}{2}}{2x^4} = 0.$$

(别犯懒! 这些都是相当简单的极限, 你现在应该试着验证它们.) 现在, 我们应用三明治定理. 由于原始函数被夹在两个 $x \to \infty$ 时都趋于 0 的函数之间, 因而它也趋于 0. 也就是说,

$$\lim_{x \to \infty} \frac{x\sin(11x^7) - \frac{1}{2}}{2x^4} = 0.$$

由不等式 $-1 \leqslant \sin(x) \leqslant 1$ (以及 $\cos(x)$ 的类似不等式) 可得到的另一个结论是, 你可以将 sin (任何东西) 或 cos (任何东西) 看作比 x 的任意正次幂次数要低, 只要你仅是在加上或减去它们. 更确切地说, 如果要求解形如

$$\lim_{x\to\infty}\frac{p(x)}{q(x)}$$

的问题, 其中 p 和 q 是多项式或多项式型函数, 但又带有一些附加的正弦和余弦, 那么即使没有这些正弦或余弦. 分子和分母的次数也不会改变. 唯一的例外是, 当 p 或 q 的次数为 0; 那样的话, 三角函数的部分将举足轻重.

以下是一个例子, 来看看附加的正弦和余弦为什么不会造成什么影响：极限

$$\lim_{x\to\infty}\frac{3x^2+2x+5+\sin(3000x^9)}{2x^2-1-\cos(22x)}$$

是什么？在分子中, $3x^2$ 仍占据主导, 因为 $\sin\left(3000x^9\right)$ 这一项只是在 -1 和 1 之间, 相较而言是无关紧要的. 与之形成对比的是, 在上一个例子中, 我们用 $\sin\left(11x^7\right)$ 和 x 的最高次数项相乘; 那里的正弦因子因而非常重要. 而在当前的例子中, 正弦项是被加上的.

分母又怎样呢？其实, 余弦项比主导项 $2x^2$ 小很多. 因此, 我们用 $3x^2$ 和分子相乘并相除, 用 $2x^2$ 和分母相乘并相除, 得到

$$\begin{aligned}\lim_{x\to\infty}\frac{3x^2+2x+5+\sin(3000x^9)}{2x^2-1-\cos(22x)} &= \lim_{x\to\infty}\frac{\dfrac{3x^2+2x+5+\sin(3000x^9)}{3x^2}\times(3x^2)}{\dfrac{2x^2-1-\cos(22x)}{2x^2}\times(2x^2)}\\ &= \lim_{x\to\infty}\frac{1+\dfrac{2}{3x}+\dfrac{5}{3x^2}+\dfrac{\sin(3000x^9)}{3x^2}}{1-\dfrac{1}{2x^2}-\dfrac{\cos(22x)}{2x^2}}\times\frac{3x^2}{2x^2}.\end{aligned}$$

现在来看看会发生什么？显然我们知道, $2/3x$、$5/3x^2$ 及 $1/2x^2$ 的极限都趋于 0, 但 $\sin\left(3000x^9\right)/3x^2$ 和 $\cos(22x)/2x^2$ 这两项会怎样呢？如果你想给出一个完整的解, 需要使用三明治定理 (对每一项使用一次) 来证明它们都趋于 0. 我建议你现在作为练习试一下. 在实践中, 大多数数学家会直接写出结果是 0, 因为我们可以确立一个一般原理：对于任意的正指数 α,

$$\lim_{x\to\infty}\frac{\sin(\text{任何东西})}{x^\alpha}=0.$$

如果用余弦替换正弦, 也会得到类似的结果. 因此, 以上极限就是

$$\frac{1+0+0+0}{1-0-0}\times\frac{3}{2}=\frac{3}{2}.$$

最后, 我们回到本章开始部分提及的例子

$$\lim_{x\to 0} x\sin\left(\frac{5}{x}\right).$$

正如我们看到的, 尽管极限是在 $x \to 0$ 时取的, 但这个极限确实属于大数的情况, 因为当 x 接近于 0 时, $5/x$ 是一个非常大的数 (正的或负的). 因此, 我们所能做的就是, 结合任何数的正弦在 -1 和 1 之间这一事实, 使用三明治定理. 特别是, 对于任意的 x, 我们有

$$-1 \leqslant \sin\left(\frac{5}{x}\right) \leqslant 1.$$

现在, 有人可能会想用 x 和以上不等式相乘, 得到

$$-x \leqslant x\sin\left(\frac{5}{x}\right) \leqslant x.$$

但不幸的是, 这只有当 $x > 0$ 时成立. 例如, 如果 $x = -2$, 那么不等式的最左边会变成 2, 最右边会变成 -2, 这简直是疯了. 因此, 先来考虑一下右极限:

$$\lim_{x\to 0^+} x\sin\left(\frac{5}{x}\right).$$

现在, 可以使用以上不等式并注意到, 当 $x \to 0^+$ 时, $-x$ 和 x 都趋于 0, 因此三明治定理适用, 上述极限是 0. 至于左极限 (当 $x \to 0^-$ 时), 我们以相同的 $\sin\left(\frac{5}{x}\right)$ 的不等式出发, 并将之乘以 x. 但这一次, 由于 x 是负的, 必须反转不等号. 特别是, 当 $x < 0$ 时, 有

$$-x \geqslant x\sin\left(\frac{5}{x}\right) \geqslant x.$$

不过这也没有太多区别. 当 $x \to 0^-$ 时, 外层的两个量仍然趋于 0, 因此中间的量也趋于 0. 由于左极限和右极限都是 0, 故双侧极限也是 0; 我们证明了

$$\lim_{x\to 0} x\sin\left(\frac{5}{x}\right) = 0.$$

(这个例子与 3.6 节的例子非常相似.)

7.1.4 “其他的” 情况



考虑极限

$$\lim_{x\to\pi/2} \frac{\cos(x)}{x - \dfrac{\pi}{2}}.$$

这次的三角函数是余弦, 且要在 $\pi/2$ 的附近求值. 这既不是小数的情况也不是大数的情况, 因此很明显, 之前的情况都不适用. 如果你只是将 $x = \pi/2$ 代入上式的话, 会得到 $0/0$ 的不定式, 这可真要命. 不过要是你了解三角函数的性质的话, 有时你也会绝境逢生. 下面就是原因.

面对 $x \to a$ 的极限, 而 $a \neq 0$ 时, 有一个很好的一般原则, 那就是**用 $t = x - a$ 作替换, 将问题转化为 $t\to 0$**. 因此, 在以上极限中, 设 $t = x - \pi/2$. 当 $x \to \pi/2$ 时,

你可以看到 $t \to 0$. 由于 $x = t + \pi/2$, 则有

$$\lim_{x\to\pi/2} \frac{\cos(x)}{x - \frac{\pi}{2}} = \lim_{t\to 0} \frac{\cos\left(t + \frac{\pi}{2}\right)}{t}.$$

注意到我们仍然需要知道余弦在 $\pi/2$ 附近的行为 (通过设 t 接近于 0, 可以看到你要取的是什么的余弦). 替换并没有改变这一事实. 这里需要你想起 2.4 节中的三角恒等式

$$\cos\left(\frac{\pi}{2} - x\right) = \sin(x).$$

在我们的极限中, 有 $\cos\left(\frac{\pi}{2} + t\right)$, 因此需要将上述三角恒等式用 $-t$ 替换 x, 得到

$$\cos\left(\frac{\pi}{2} + t\right) = \sin(-t).$$

我们还需要记得, 正弦函数是一个奇函数. 因此, 事实上,

$$\cos\left(\frac{\pi}{2} + t\right) = \sin(-t) = -\sin(t).$$

现在, 可以将它代入极限并完成问题的求解. 也就是说,

$$\lim_{x\to\pi/2} \frac{\cos(x)}{x - \frac{\pi}{2}} = \lim_{t\to 0} \frac{\cos\left(t + \frac{\pi}{2}\right)}{t} = \lim_{t\to 0} \frac{-\sin(t)}{t} = -1.$$

并没有那么轻松惬意, 但了解三角恒等式显然会对求解类似的情况有所帮助.

7.1.5 一个重要极限的证明

本章反复使用了以下极限, 现在是时候证明它了：

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1.$$

证明显然需要借助直角三角形的几何学, 因为那也正是正弦函数诞生的地方. 让我们先从右极限 $(x \to 0^+)$ 开始. 一旦做到了这一步, 我们会发现双侧极限其实相当简单. 因此, 先假设 x 接近于 0 但为正. 让我们来画一个以 O 为中心、夹角为 x、半径为 1 的扇形 OAB, 如图 7-2 所示. 我们将对这幅图进行一些操作, 但首先有一个问题：这个扇形的面积是什么呢？

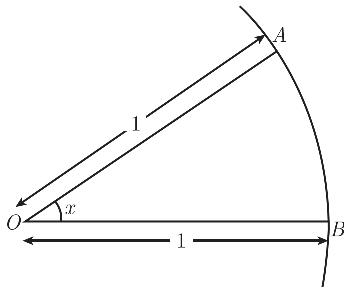


图 7-2

不妨想象这个扇形是一整块比萨中的一片. 这块比萨的半径为 1 个单位, 因此它的面积是 $\pi r^2 = \pi$ 平方单位. 现在, 这一片占整块比萨的多少呢？整块比萨有 2π 弧度的角, 而这一片的夹角是 x, 因此这一片占整块比萨的 $x/2\pi$. 故其面积为 $(x/2\pi)\times\pi$, 即 $x/2$ 平方单位. 也就是说,

$$\text{扇形 } OAB \text{ 的面积} = \frac{x}{2} \text{ 平方单位.}$$

(这里有一个一般公式：夹角为 x 弧度、半径为 r 个单位的扇形的面积就是 $xr^2/2$ 平方单位.)

现在, 在这幅图上进行一些操作. 首先, 连接 AB 画一条直线. 然后, 从 A 出发画一条垂线到直线 OB, 称 C 为基点. 还要将直线 OA 向外延长一些, 最后画出圆在点 B 的切线. 这条切线和延长的直线 OA 相交于点 D. 在完成了上述操作之后, 我们得到图 7-3. 我在图中标记了 AC 和 DB 的长度. 要看出我是如何算出它们的, 需要注意到 $\sin(x) = \frac{|AC|}{|OA|}$ (记住, $|AC|$ 表示 "线段 AC 的长度"). 由于 $|OA| = 1$, 我们有 $|AC| = \sin(x)$. 同样, 有 $\tan(x) = \frac{|DB|}{|OB|}$, 由于 $|OB| = 1$, 因此 $|DB| = \tan(x)$.

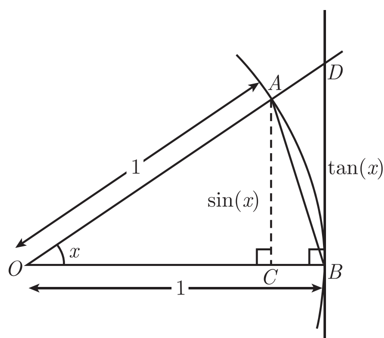


图　7-3

我想将精力集中在三个对象上. 一是原始的扇形, 我们已经求出它的面积是 $x/2$ 平方单位. 我们再来看看 $\triangle OAB$ 和 $\triangle OBD$. $\triangle OAB$ 的底是 OB, 其长度为 1 个单位, 高是 AC, 其长度是 $\sin(x)$ 个单位. 因此, $\triangle OAB$ 的面积是底乘高的一半, 即 $\sin(x)/2$ 平方单位. 至于 $\triangle OBD$, 它的底是 OB, 其长度是 1 个单位, 它的高是 DB, 其长度是 $\tan(x)$ 个单位. 因此, $\triangle OBD$ 的面积是 $\tan(x)/2$ 平方单位. 这里关键的一个观察：

$\triangle OAB$ 包含在扇形 OAB 中, 扇形 OAB 又包含在 $\triangle OBD$ 中.

这意味着, $\triangle OAB$ 的面积小于扇形 OAB 的面积, 扇形 OAB 的面积又小于 $\triangle OBD$ 的面积：

$$\triangle OAB \text{ 的面积} < \text{扇形 } OAB \text{ 的面积} < \triangle OBD \text{ 的面积.}$$

我们知道所有这三个量可以用变量 x 表示; 将它们代入, 有

$$\frac{\sin(x)}{2} < \frac{x}{2} < \frac{\tan(x)}{2}.$$

用 2 和以上不等式相乘, 我们会得到一个非常好的值得记忆的不等式：

$$\boxed{\sin(x) < x < \tan(x), \ 0 < x < \frac{\pi}{2}.}$$

现在可以求极限了. 首先来取这个不等式的倒数. 要记住, 这个操作将使我们把小于号变为大于号. 我们写出 $\tan(x) = \sin(x) / \cos(x)$, 不等式的倒数就是

$$\frac{1}{\sin(x)} > \frac{1}{x} > \frac{\cos(x)}{\sin(x)}.$$

最后, 用正的量 $\sin(x)$ 和上式相乘, 得到

$$1 > \frac{\sin(x)}{x} > \cos(x).$$

如果你感到这个顺序不舒服, 它总是可以被重写为

$$\cos(x) < \frac{\sin(x)}{x} < 1.$$

(要记住, 这对于任意的介于 0 和 $\pi/2$ 的 x 都成立.) 现在, 使用三明治定理：由于 $\cos(0) = 1$ 且 $y = \cos(x)$ 是连续的, 我们知道 $\lim\limits_{x\to 0^+} \cos(x) = 1$; 同样, $\lim\limits_{x\to 0^+} (1) = 1$; 而量 $\sin(x)/x$ 被夹在 $\cos(x)$ 和 1 之间, 当 $x \to 0^+$ 时, 后两者都趋于 1. 因此根据三明治定理,

$$\lim_{x\to 0^+} \frac{\sin(x)}{x} = 1.$$

这样就求出了右极限.

我们仍然需要处理左极限并证明

$$\lim_{x\to 0^-} \frac{\sin(x)}{x} = 1.$$

如果可以做到的话, 那么就证明了左极限与右极限均为 1, 因此双侧极限也是 1, 这样就完成了证明.

为了证明左极限是 1, 我们设 $t = -x$. 那么当 x 是一个很小的负数时, t 是一个很小的正数. 用数学符号表达可以说, 当 $x \to 0^-$ 时, 有 $t \to 0^+$. 因此, 以上极限可以写作

$$\lim_{t\to 0^+} \frac{\sin(-t)}{-t}.$$

由于 $\sin(-t) = -\sin(t)$ (因为正弦函数是奇函数), 可以将以上极限简化为

$$\lim_{t\to 0^+} \frac{-\sin(t)}{-t} = \lim_{t\to 0^+} \frac{\sin(t)}{t}.$$

我们已经看到该极限是 1, (好吧, 前面是 x 而不是 t, 但这又有什么关系呢？) 这样就完成了证明.

在继续讨论三角函数求导之前, 我想先考虑一下 $f(x) = \sin(x)/x$ 的图像. 左极限的论证过程其实也证明了 f 是一个偶函数. (你能看出来吗?) 这意味着, y 轴就像是 $y = f(x)$ 的图像的一面镜子. 如果回顾一下 3.5 节的内容, 你可以看到已经画出了当 $x > 3$ 时的 $y = f(x)$ 的图像. 没有画 $x \leqslant 3$ 的图像是因为我们不知道那里会发生什么. 现在我们知道了：当 $x \to 0$ 时, 量 $f(x) = \sin(x)/x \to 1$. 事实上, 我们证明了 $\sin(x)/x$ 位于 $\cos(x)$ 和 1 之间. 这样可以将图像扩展到 $x > 0$. 最后, 我们使用 f 的偶函数性来画出 $y = \sin(x)/x$ 的完整的图像 (注意到 x 轴和 y 轴的比例不同), 如图 7-4 所示.

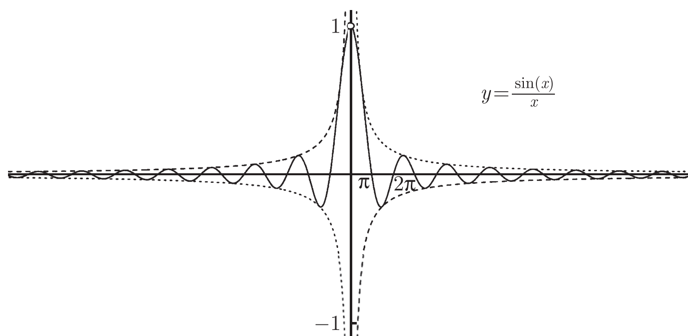


图　7-4

包络函数 $y=1/x$ 和 $y=-1/x$ 的图像用虚线表示. 此外, x 轴截距是所有的除 0 之外的 π 的倍数. 最后, 正如你看到的, 该函数在 $x=0$ 上不连续, 因为它在那里无定义. 然而, 如果我们定义函数 g, 使其定义为, 如果 $x \neq 0, g(x)=\sin(x)/x$ 及 $g(0)=1$, 那么实际上就填充了图中在 $(0,1)$ 处的空心圆, 并且函数 g 是连续的.

7.2　三角函数的导数

现在是时候对某些三角函数求导了. 让我们先对 $\sin(x)$ 关于 x 求导. 为了做到这点, 将要使用 7.1.2 节中的两个极限：

$$\lim_{h\to 0}\frac{\sin(h)}{h}=1 \quad 和 \quad \lim_{h\to 0}\frac{1-\cos(h)}{h}=0.$$

(好吧, 我将 x 变为了 h, 但这没关系 —— h 是一个虚拟变量, 它可以被任意字母所替换.) 不管怎样, 令 $f(x)=\sin(x)$, 让我们对其求导：

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\lim_{h\to 0}\frac{\sin(x+h)-\sin(x)}{h}.$$

现在该怎么办呢? 其实, 你应该回想一下公式

$$\sin(A+B)=\sin(A)\cos(B)+\cos(A)\sin(B);$$

如果想不起来, 最好再看一下第 2 章. 不管怎样, 我们想要用 x 替换 A, 用 h 替换 B, 因此有

$$\sin(x+h)=\sin(x)\cos(h)+\cos(x)\sin(h).$$

将上式代入以上极限, 会得到

$$f'(x)=\lim_{h\to 0}\frac{\sin(x)\cos(h)+\cos(x)\sin(h)-\sin(x)}{h}.$$

最后剩下的工作就是将这些项进行重组和提取公因子, 得到

$$f'(x) = \lim_{h \to 0} \frac{\sin(x)(\cos(h) - 1) + \cos(x)\sin(h)}{h}$$
$$= \lim_{h \to 0} \left(\sin(x) \left(\frac{\cos(h) - 1}{h} \right) + \cos(x) \left(\frac{\sin(h)}{h} \right) \right).$$

注意到我们将含有 x 的部分与含有 h 的部分尽可能地分离开. 现在, 实际上要求当 $h \to 0$(不是 $x \to 0$!) 时的极限. 使用本节开始部分的两个极限, 会得到

$$f'(x) = \sin(x) \times 0 + \cos(x) \times 1 = \cos(x).$$

也就是说, $f(x) = \sin(x)$ 的导数是 $f'(x) = \cos(x)$, 或者换句话说,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \cos(x).}$$

现在, 你应该试着用 $f(x) = \cos(x)$ 重复一下以上论证. 你只需要用到第 2 章的恒等式

$$\cos(A + B) = \cos(A)\cos(B) - \sin(A)\sin(B).$$



这是个很好的练习, 因此现在就尝试一下吧. 如果你求解正确, 应该会看到

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = -\sin(x).}$$

不管怎样, 现在去求其他三角函数的导数就是小菜一碟了. 你不需要使用任何极限, 可以只使用商法则和链式求导法则. 让我们从 $y = \tan(x)$ 的导数开始. 可以将 $\tan(x)$ 写成 $\sin(x)/\cos(x)$, 因此如果设 $u = \sin(x)$ 及 $v = \cos(x)$, 那么 $y = u/v$. 我们已经求出 $\mathrm{d}u/\mathrm{d}x = \cos(x)$ 及 $\mathrm{d}v/\mathrm{d}x = -\sin(x)$, 所以使用商法则, 可以得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2} = \frac{\cos(x)(\cos(x)) - \sin(x)(-\sin(x))}{\cos^2(x)}.$$

最后一个分式的分子就是 $\cos^2(x) + \sin^2(x)$, 它总是等于 1. 因此, 导数就是

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\cos^2(x)} = \sec^2(x).$$

这样就证明了

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\tan(x) = \sec^2(x).}$$

现在, 来计算 $y = \sec(x)$ 的导数. 这里可以写成 $y = 1/\cos(x)$, 所以或许你会认为使用商法则是最好的. 确实, 你可以使用商法则, 但链式求导法则其实更好一些. 如果 $u = \cos(x)$, 那么 $y = 1/u$. 我们可以对这两个函数求导：$\mathrm{d}y/\mathrm{d}u = -1/u^2$ 及 $\mathrm{d}u/\mathrm{d}x = -\sin(x)$. 根据链式求导法则,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = \left(-\frac{1}{u^2}\right)(-\sin(x)) = \frac{\sin(x)}{\cos^2(x)}.$$

在上面, 最后必须用 $\cos(x)$ 替换 u. 事实上, 可以再整理一下, 得到

$$\frac{\sin(x)}{\cos^2(x)} = \frac{1}{\cos(x)}\frac{\sin(x)}{\cos(x)} = \sec(x)\tan(x),$$

这样就证明了

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\sec(x) = \sec(x)\tan(x).}$$

至于 $y = \csc(x)$, 它应该被写成 $1/\sin(x)$. 再一次地, 我们最好使用链式求导法则, 令 $u = \sin(x)$ 及 $y = 1/u$. 但我知道你禁不住想要使用商法则, 因为它是一个商的形式, 尽管这其实是下策. 你就是不相信我. 那好吧, 我们来看一下. 要对 $y = 1/\sin(x)$ 使用商法则, 实际上要令 $u = 1$ 及 $v = \sin(x)$. 那么 $\mathrm{d}u/\mathrm{d}x = 0$ 及 $\mathrm{d}v/\mathrm{d}x = \cos(x)$. 根据商法则,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2} = \frac{\sin(x)(0) - 1(\cos(x))}{\sin^2(x)} = -\frac{\cos(x)}{\sin^2(x)}.$$

好吧, 这也没有那么糟糕, 但使用链式求导法则仍然会更好些. 不管怎样, 正如刚刚对 $y = \sec(x)$ 所做的, 我们可以整理得到

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\csc(x) = -\csc(x)\cot(x).}$$

最后, 考虑 $y = \cot(x)$, 而它当然可以被写成 $y = \cos(x)/\sin(x)$ 或 $y = 1/\tan(x)$. 你可以在 $y = \cos(x)/\sin(x)$ 上使用商法则, 又或者既然已经知道 $\tan(x)$ 的导数, 你也可以在 $y = 1/\tan(x)$ 上使用链式求导法则 (或商法则). 你甚至还可以将 $\cot(x)$ 写成 $\cos(x)\csc(x)$ 的形式, 并使用乘积法则. 不管你用哪种方式, 应该会得到

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\cot(x) = -\csc^2(x).}$$

你应该用心记住所有的这六个方框公式. 要注意到在三个互余函数 (余弦、余割、余切) 之前都有一个负号, 并且导数是正常导数的 co– 形式. 例如, $\sec(x)$ 的导数是 $\sec(x)\tan(x)$, 因此在所有东西前面都加上一个 "co", 并再加一个负号, 我们得到 $\csc(x)$ 的导数是 $-\csc(x)\cot(x)$. 这对于余弦和余切也成立, 要知道 (对于余弦), co-co-sine 正好是原始的正弦函数.

顺便说一下, $f(x) = \sin(x)$ 的二阶导是什么呢? 我们知道 $f'(x) = \cos(x)$, 这样, $f''(x)$ 就是 $\cos(x)$ 的导数, 而这正是我们之前看到的 $-\sin(x)$. 也就是说,

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}(\sin(x)) = -\sin(x).$$

该函数的二阶导正好是负的原始函数. 这对于 $g(x) = \cos(x)$ 也成立. 这种事情就不会发生在 (非零的) 多项式上, 因为一个多项式的导数是一个次数比原始多项式的次数低一次的新的多项式.

7.2.1 求三角函数导数的例子

由于现在你需要对更多的函数求导, 牢牢记住如何使用乘积法则、商法则以及链式求导法则依旧十分重要. 例如, 如何求下列导数:

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^2\sin(x)), \quad \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\sec(x)}{x^5}\right) \quad 和 \quad \frac{\mathrm{d}}{\mathrm{d}x}(\cot(x^3))?$$

让我们一个一个地求解吧. 如果 $y = x^2\sin(x)$, 那么可以写出 $y = uv$, 其中 $u = x^2$ 及 $v = \sin(x)$. 现在, 我们需要整理出那张表:

$$\begin{aligned} u &= x^2 & v &= \sin(x) \\ \frac{\mathrm{d}u}{\mathrm{d}x} &= 2x & \frac{\mathrm{d}v}{\mathrm{d}x} &= \cos(x). \end{aligned}$$

使用乘积法则 (参见 6.2.3 节), 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x} = \sin(x)\cdot(2x) + x^2\cos(x).$$

这通常会被写成 $2x\sin(x) + x^2\cos(x)$. 不管怎样, 让我们接着来做第二个例子. 如果 $y = \sec(x)/x^5$, 这一次设 $u = \sec(x)$ 及 $v = x^5$, 这样 $y = u/v$. 我们的表如下:

$$\begin{aligned} u &= \sec(x) & v &= x^5 \\ \frac{\mathrm{d}u}{\mathrm{d}x} &= \sec(x)\tan(x) & \frac{\mathrm{d}v}{\mathrm{d}x} &= 5x^4. \end{aligned}$$

使用商法则, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v\dfrac{\mathrm{d}u}{\mathrm{d}x} - u\dfrac{\mathrm{d}v}{\mathrm{d}x}}{v^2} = \frac{x^5\sec(x)\tan(x) - \sec(x)\cdot 5x^4}{(x^5)^2} = \frac{\sec(x)(x\tan(x) - 5)}{x^6}.$$

注意到最后消去了因子 x^4. 现在来看看第三个例子. 设 $y = \cot\left(x^3\right)$. 这里我们正在处理两个函数的复合, 因此最好使用链式求导法则. 首先在 x 上进行的操作是立方, 因此令 $u = x^3$, 那么 $y = \cot(u)$. 汇总如下:

$$\begin{aligned} y &= \cot(u) & u &= x^3 \\ \frac{\mathrm{d}y}{\mathrm{d}u} &= -\csc^2(u) & \frac{\mathrm{d}u}{\mathrm{d}x} &= 3x^2. \end{aligned}$$

根据链式求导法则, 我们有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = -\csc^2(u)\cdot 3x^2.$$

我们不能把项 u 留在那里, 需要用 x^3 替换它. 因此, 要求的导数就是 $-3x^2\csc^2(x^3)$.



在继续前进之前, 我想告诉你一个小诀窍. 假设你有 $y=\sin(8x)$, 并且想要求 $\mathrm{d}y/\mathrm{d}x$. 你应该使用链式求导法则, 设 $u=8x$, 这样 $y=\sin(u)$. 很容易得出 $\mathrm{d}y/\mathrm{d}x=8\cos(8x)$ (试着做一下!). 当然, 数 8 没什么特别的, 它可以是任何数. 因此, 有一个一般法则：对于任意的常数 a,

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sin(ax))=a\cos(ax).$$



基本上, **如果用 ax 替换 x, 那么当求导的时候, 在最前面会有一个额外的因子 a**. 这对于其他三角函数也适用. 例如, $\tan(x)$ 关于 x 的导数是 $\sec^2(x)$, 因此 $\tan(2x)$ 的导数是 $2\sec^2(2x)$. 同样地, $\csc(x)$ 的导数是 $-\csc(x)\cot(x)$, 因此 $\csc(19x)$ 的导数是 $-19\csc(19x)\cot(19x)$. 这可以让你在这些简单的情况下免去使用链式求导法则.

7.2.2　简谐运动

三角函数自然而然会出现的一个地方是描述弹簧振子的运动. 事实表明, 如果 x 是弹簧振子在时刻 t 的位置, 我们取向上的方向作为正方向, 那么描述 x 的方程大致类似于 $x=3\sin(4t)$. 数 3 和 4 可能会改变, “正弦” 也可能变成 “余弦”, 但基本思想就是那样. 该方程是合理的. 毕竟, 余弦函数总是来回振荡, 而振子也是如此. 这种类型的运动被称为**简谐运动**.

因此, 如果 $x=3\sin(4t)$ 是振子从初始点出发的位移, 那么在时刻 t 振子的速度和加速度是多大呢？需要做的就是求导. 我们知道 $v=\mathrm{d}x/\mathrm{d}t$, 因此只需要对 $3\sin(4t)$ 关于 t 求导. 可以使用链式求导法则, 但使用上一节结尾的那个结论会更简单. 确实, 为了对 $\sin(4t)$ 关于 t 求导, 我们观察到 $\sin(t)$ 的导数是 $\cos(t)$, 因此 $\sin(4t)$ 的导数就是 $4\cos(4t)$. (不要忘记在最前面有一个 4!) 总而言之, 我们有

$$v=\frac{\mathrm{d}}{\mathrm{d}t}(3\sin(4t))=3\times 4\cos(4t)=12\cos(4t).$$

现在可以对加速度, 它由 $\mathrm{d}v/\mathrm{d}t$ 给出, 重复这个过程. 使用相同的技巧, 得到

$$a=\frac{\mathrm{d}v}{\mathrm{d}t}=\frac{\mathrm{d}}{\mathrm{d}t}(12\cos(4t))=-12\times 4\sin(4t)=-48\sin(4t).$$

注意到加速度 (当然它就是位移的二阶导) 基本上和位移本身是一样的, 除了最前面有一个负号以及系数有所不同 (48 取代了 3). 这个负号表示加速度和位移的方向是相反的. 事实上, 由于 $48=3\times 16$, 这就证明了

$$a=-16x.$$

现在, 为了阐释这个方程, 让我们更深入地研究一下振子的运动.

位置 x 由 $x = 3\sin(4t)$ 给出, 当振子在平衡位置时有 $x = 0$. 现在, 如果用 3 和不等式 $-1 \leqslant \sin(4t) \leqslant 1$ 相乘 (这对所有的 t 都成立), 我们得到 $-3 \leqslant 3\sin(4t) \leqslant 3$. 也就是说, $-3 \leqslant x \leqslant 3$. 因此, 可以看到 x 在 -3 和 3 之间振荡. 当 x 为正时, 振子在平衡位置的上方, 那么 a 是负的, 这很好：加速度是向下的, 一如它理当如此. 当 x 变得越来越大时, 弹簧压缩得更厉害, 致使振子经受一个更大的力和向下的加速度. 最终, 振子开始向下运动, 不一会儿, x 变为负的. 然后, 振子在它平衡位置的下方, 因此弹簧被伸展并要将振子拉回来. 确实, 当 x 为负时, a 为正, 因此力是向上的. 图 7-5 显示了整个过程.

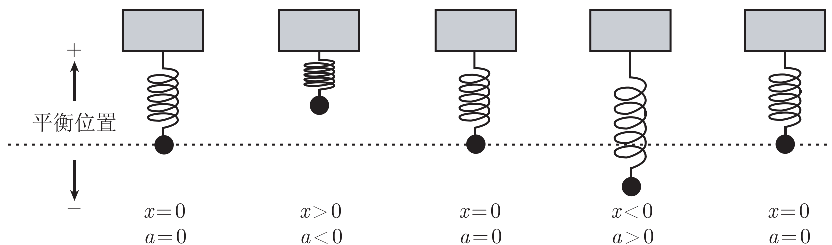


图 7-5

当振子在它运动的最上方时, 其速度为 0. 由于有 $v = 12\cos(4t)$, 当 $4t$ 是 $\pi/2$ 的奇数倍, 即 $t = (2n+1)\pi/8$ (n 为某个整数) 时, 这个情况会出现. 现在, 说够了简谐运动, 我们再多看一个三角函数求导的例子, 然后就进入下一章的隐函数求导.

7.2.3 一个有趣的函数

试考虑函数 f, 其定义为

$$f(x) = x^2 \sin\left(\frac{1}{x}\right).$$

它的导数是什么呢？我们不用担心 $x = 0$, 因为 f 在那里无定义, x 的其他值则都没有问题. 设 $y = f(x)$, 那么 y 是 $u = x^2$ 和 $v = \sin(1/x)$ 的乘积. 对 u 关于 x 求导很简单 (结果就是 $2x$), 但 v 有点难. 最好的猜测是设 $w = 1/x$, 这样 $v = \sin(w)$. 然后, 可以给出那张标准表：

$$\begin{array}{cc} v = \sin(w) & w = \dfrac{1}{x} \\ \dfrac{dv}{dw} = \cos(w) & \dfrac{dw}{dx} = -\dfrac{1}{x^2}. \end{array}$$

现在, 可以使用链式求导法则：

$$\frac{dv}{dx} = \frac{dv}{dw}\frac{dw}{dx} = \cos(w)\left(-\frac{1}{x^2}\right) = -\frac{\cos(1/x)}{x^2}.$$

由于有了 du/dx 和 dv/dx, 我们最后得以在 $y = uv$ 上使用乘积法则：

$$\frac{\mathrm{d}y}{\mathrm{d}x} = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x} = \sin\left(\frac{1}{x}\right)(2x) + x^2\left(-\frac{\cos(1/x)}{x^2}\right) = 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

这样就完成了求解.



事实表明, 函数 f 相当有趣. 我们来看看为什么. (如果你没有这种感觉, 我猜你大可先跳到下一章, 以后再回来看.) 不管怎样, 为了进一步研究, 需要以下三个极限：

$$\lim_{x\to 0} x^2\sin\left(\frac{1}{x}\right) = 0, \quad \lim_{x\to 0} x\sin\left(\frac{1}{x}\right) = 0 \quad \text{和} \quad \lim_{x\to 0^+}\cos\left(\frac{1}{x}\right) \text{ 不存在.}$$

你可以使用三明治定理以及任何东西 (甚至是 $1/x$) 的正弦或余弦都在 -1 和 1 之间的这一事实来求解这三个极限中的前两个. 第三个极限稍微复杂些, 但我们在 3.3 节中已经讨论过 $\sin(1/x)$, 而将正弦改为余弦并没有什么区别. 问题 (你可能还记得) 是, 当 $x \to 0^+$ 时, $\cos(1/x)$ 在 -1 和 1 之间的振荡变得越来越激烈, 因此极限不存在.

不管怎样, 第一个极限是说 $\lim\limits_{x\to 0} f(x) = 0$, 尽管 $f(0)$ 是无定义的. 这意味着, 通过填充点 $f(0) = 0$, 可以将 f 扩展为连续函数. 因此, 我们抛弃旧的 f 并由以下公式定义一个新的 f：

$$f(x)\begin{cases} x^2\sin\left(\dfrac{1}{x}\right) & \text{如果 } x \neq 0, \\ 0 & \text{如果 } x = 0. \end{cases}$$

我们刚刚证明了这个改善后的 f 是处处连续的. 我们已经求出当 $x \neq 0$ 时它的导数是

$$f'(x) = 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

那么在 $x = 0$ 处 f 的导数又是什么呢? 在这里没有一个法则能够帮得上忙, 我们必须使用导数的定义公式：

$$f'(0) = \lim_{h\to 0}\frac{f(0+h) - f(0)}{h} = \lim_{h\to 0}\frac{h^2\sin(1/h) - 0}{h} = \lim_{h\to 0} h\sin\left(\frac{1}{h}\right).$$

最后这个极限就是之前三个极限中的中间那个 (用 h 替换 x), 这个极限存在且值为 0. 这意味着 f 实际上在 $x = 0$ 处可导. 事实上, $f'(0) = 0$. 从 $y = f(x)$ 时图像上你能看出这点吗? 图 7-6 就是 $-0.1 < x < 0.1$ 时图像的样子, 伴有包络函数 $y = x^2$ 和 $y = -x^2$.

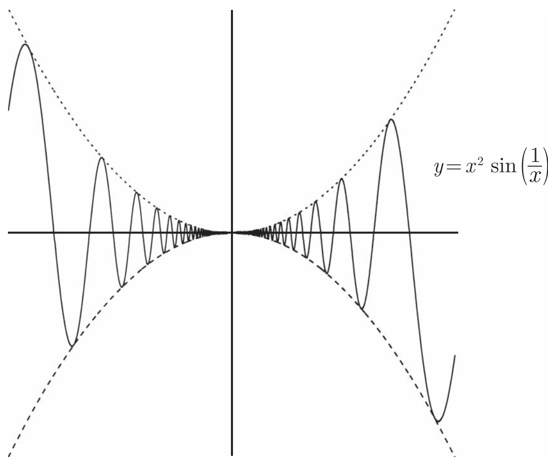


图 7-6

在 $x=0$ 上, 它在我看来很不稳定, 似乎根本无法确认在那里导数会存在 —— 但刚刚证明了它存在! 这就引出以下问题：

$$\lim_{x\to 0^+} f'(x)$$

是什么呢? 由于我们知道 $f'(0)=0$, 你或许会认为以上极限就是 0. 利用前述 $f'(x)$ 在 $x\neq 0$ 时的公式, 让我们来检验一下吧：

$$\lim_{x\to 0^+} f'(x) = \lim_{x\to 0^+}\left(2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)\right).$$

这里有两项需要处理. 第一项 $(2x\sin(1/x))$ 的极限趋于 0, 因为它就是之前三个极限中的中间那个的两倍. 另一方面, 第二项 $(\cos(1/x))$ 当 $x\to 0$ 时极限不存在, 这正是前述第三个极限所表达的意思. 因此结论是, $\lim\limits_{x\to 0^+} f'(x)$ 不存在. 根据对称性 (检验一下 f 是一个奇函数), $\lim\limits_{x\to 0^-} f'(x)$ 也不存在.

现在来总结一下我们的发现. 我们的函数 f 处处连续且处处可导, 甚至在 $x=0$ 处也不例外. 事实上, 在 $x=0$ 处, 导数 $f'(0)$ 等于 0, 但在 0 的附近, 导数 $f'(x)$ 振荡得很激烈：$\lim\limits_{x\to 0} f'(x)$ 不存在, 尽管 $f'(0)$ 存在. 特别是, 我们证明了导函数 f' 本身不是连续函数. 因此, 存在本身可导但其导数不连续的函数. 这真是十分的有趣!

第 8 章　隐函数求导和相关变化率

过去几章, 我们都在试图对眼前的一切进行求导. 现在暂且打断一下, 是时候来看一下隐函数的求导, 后者是对常规求导的一个很好的一般化. 之后, 我们会看到如何使用这种技巧来求解涉及变化的量的应用问题. 如果知道一个量的变化有多快, 我们就能求出另一个不同的但与之相关的量的变化会有多快. 总之, 本章的主要内容正如标题所示:

- 隐函数求导;
- 相关变化率.

8.1　隐函数求导

考虑两个导数:

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^2) \quad 和 \quad \frac{\mathrm{d}}{\mathrm{d}x}(y^2).$$

正如我们已经看到的, 第一个就是 $2x$. 那么第二个是 $2y$ 吗? 如果是关于 y 求导, 那么结果就是它, 但这不是关于 y 求导, 在分母上的 $\mathrm{d}x$ 告诉我们这是在关于 x 求导. 我们该如何处理呢?

最好的方法是告诉自己, 上面第一个导数问的是: 当对 x 稍作改变时, 量 x^2 会有多大变化. 正如我们在 5.2.7 节中看到的, 如果对 x 稍作改变, 那么 x^2 就会有近似 $2x$ 倍那么多的变化.

另一方面, 如果对 x 稍作改变, y^2 会怎么样? 这正是为了求出上面第二个导数 $\mathrm{d}\left(y^2\right)/\mathrm{d}x$ 所需要知道的. 不妨这样思考: 如果改变 x, 那么 y 会有点变化; y 的这个变化又会引起 y^2 的变化. (当然, 这一切只有当 y 依赖于 x 时才正确. 如果不是这样的话, 那么改变 x 时, y 根本不会有任何变化.)

如果你认为这听起来我好像是在暗示链式求导法则, 那你想得一点没错. 以下就是它如何具体动作. 令 $u = y^2$, 则 $\mathrm{d}u/\mathrm{d}y = 2y$. 根据链式求导法则,

$$\frac{\mathrm{d}}{\mathrm{d}x}(y^2) = \frac{\mathrm{d}u}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = 2y\frac{\mathrm{d}y}{\mathrm{d}x}.$$

因此, 如果对 x 稍作改变, 那么 y^2 会有 $2y\,(\mathrm{d}y/\mathrm{d}x)$ 倍的变化. 现在你或许会抱怨结果中还是包含 $\mathrm{d}y/\mathrm{d}x$, 但又能怎么办呢? 如果你想要知道当对 x 稍作改变时, 量 y^2 会如何变化, 那你势必首先需要知道 y 是如何变化的! (此外, 如果 y 不依赖于 x, 那么对于所有的 x, $\mathrm{d}y/\mathrm{d}x$ 都等于 0, 故对于所有的 x, $\mathrm{d}\left(y^2\right)/\mathrm{d}x$ 也是 0. 也就是

说, y^2 也不依赖于 x.)

8.1.1 技巧和例子



现在该来实际应用一下了. 考虑方程

$$x^2 + y^2 = 4.$$

这里量 y 不是 x 的函数. 事实上, 当 $-2 < x < 2$, 有两个 y 值满足这个方程. 另一方面, 上述关系的图像就是半径为 2、圆心位于原点的单位圆. 该圆处处有切线, 并且不用写出 $y = \pm\sqrt{4 - x^2}$ 并求导, 我们就应该能够求出它们的斜率. 事实上, 所要做的只是在等号两边添加一个 $\mathrm{d}/\mathrm{d}x$:

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^2 + y^2) = \frac{\mathrm{d}}{\mathrm{d}x}(4).$$

正如我们所知, 等号左边可以直接拆分成两部分. 事实上, 通常可以直接写出

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^2) + \frac{\mathrm{d}}{\mathrm{d}x}(y^2) = \frac{\mathrm{d}}{\mathrm{d}x}(4).$$

为了化简上式, 注意到我们已经在上一节求出了左边的两个量, 而右边的为零, 因为 4 是常数. 当心不要写成 4, 这是个非常常见的错误! 不管怎么说, 我们得到

$$2x + 2y\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

将上式除以 2 并整理得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x}{y}.$$

这个公式说的是, 圆上点 (x, y) 处的切线的斜率是 $-x/y$. 如果该点不在圆上, 那么此公式没有告诉我们什么东西 (至少就我们所关心的而言). 现在, 我们使用公式来求圆上点 $(1, \sqrt{3})$ 处的切线方程. 该点的确位于圆上, 原因是 $x^2 + y^2 = 1^2 + \left(\sqrt{3}\right)^2 = 4$. 根据上述公式, 斜率由 $\mathrm{d}y/\mathrm{d}x = -1/\sqrt{3}$ 给出. 因此, 切线的斜率是 $-1/\sqrt{3}$ 并且通过 $(1, \sqrt{3})$. 使用点斜式公式, 可以看到直线的方程是

$$y - \sqrt{3} = -\frac{1}{\sqrt{3}}(x - 1).$$

如果你喜欢, 这可以稍加简化为 $y = (4 - x)/\sqrt{3}$.



下面是另一个例子: 如果

$$5\sin(x) + 3\sec(y) = y - x^2 + 3,$$

在原点处的切线方程是什么呢? 和先前的例子不同, 这一次不可能通过解方程求得 y(或 x). 因此, 必须使用隐函数求导. 让我们首先检验原点确实位于曲线上. 将 $x = 0$ 和 $y = 0$ 代入上式得出左边为 $5\sin(0) + 3\sec(0)$, 这正好是 3 (回想一下, $\sec(0) = 1/\cos(0) = 1$). 右边也是 3, 因此原点在曲线上. 现在, 对以上方程求导, 并将它拆分开:

$$\frac{\mathrm{d}}{\mathrm{d}x}(5\sin(x))+\frac{\mathrm{d}}{\mathrm{d}x}(3\sec(y))=\frac{\mathrm{d}y}{\mathrm{d}x}-\frac{\mathrm{d}}{\mathrm{d}x}(x^2)+\frac{\mathrm{d}}{\mathrm{d}x}(3).$$

这些量中唯一难于简化的是左边的第二项. 不过它也没有那么难：令 $u=3\sec(y)$, 那么 $\mathrm{d}u/\mathrm{d}y=3\sec(y)\tan(y)$, 于是根据链式求导法则, 有

$$\frac{\mathrm{d}}{\mathrm{d}x}(3\sec(y))=\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{\mathrm{d}u}{\mathrm{d}y}\cdot\frac{\mathrm{d}y}{\mathrm{d}x}=3\sec(y)\tan(y)\frac{\mathrm{d}y}{\mathrm{d}x}.$$

回到先前的方程并对两边求导, 于是得到

$$5\cos(x)+3\sec(y)\tan(y)\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}x}-2x.$$

注意到当对常数 3 求导时, 你会得到 0. 无论如何, 这里可以求解 $\mathrm{d}y/\mathrm{d}x$：只需将所有包含 $\mathrm{d}y/\mathrm{d}x$ 的部分移到等号的一边, 而其他的各项移到等号的另一边, 即

$$\frac{\mathrm{d}y}{\mathrm{d}x}-3\sec(y)\tan(y)\frac{\mathrm{d}y}{\mathrm{d}x}=2x+5\cos(x).$$

现在提取公因式, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x}(1-3\sec(y)\tan(y))=2x+5\cos(x).$$

然后做除法, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{2x+5\cos(x)}{1-3\sec(y)\tan(y)}.$$

最后, 将 $x=0$ 和 $y=0$ 代入上式, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{2(0)+5\cos(0)}{1-3\sec(0)\tan(0)}=\frac{2(0)+5(1)}{1-2(1)(0)}=5.$$

由于切线的斜率是 5, 并且通过原点, 其方程就是 $y=5x$, 这样就完成了求解. 但你看出了怎么做我们或许可以省点儿力气吗？回到上述方程

$$5\cos(x)+3\sec(y)\tan(y)\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}x}-2x.$$

刚才我们对此摆弄了一番, 试图求出 $\mathrm{d}y/\mathrm{d}x$ 的一般表达式, 但实际上我们只关心在原点处会发生什么. 因此, 通过将 $x=0$ 和 $y=0$ 代入以上方程, 就可以节省点儿时间. 我们会得到

$$5\cos(0)+3\sec(0)\tan(0)\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}x}-2(0).$$

这很容易简化为 $\mathrm{d}y/\mathrm{d}x=5$. 因此, 一条有用的经验法则是, **如果你需要的只是一个特定点上的导数, 不妨在重新整理之前就做替换** —— 这经常能节省时间.



到目前为止, 我们只使用了链式求导法则. 有时候, 你或许需要使用乘积法则或商法则. 例如, 如果

$$y\cot(x)=3\csc(y)+x^7,$$

那么将需要用到乘积法则和链式求导法则来求 $\mathrm{d}y/\mathrm{d}x$. 确实, 如果求导, 将得到

$$\frac{\mathrm{d}}{\mathrm{d}x}(y\cot(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(3\csc(y)) + \frac{\mathrm{d}}{\mathrm{d}x}(x^7).$$

左边是 y 和 $\cot(x)$ 的乘积. 我们应该给它一个名字, 比如 s, 这样 $s = y\cot(x)$. 如果也令 $v = \cot(x)$, 那么 $s = yv$, 进而可以使用乘积法则来对 s 关于 x 求导：

$$\frac{\mathrm{d}s}{\mathrm{d}x} = v\frac{\mathrm{d}y}{\mathrm{d}x} + y\frac{\mathrm{d}v}{\mathrm{d}x} = \cot(x)\frac{\mathrm{d}y}{\mathrm{d}x} + y(-\csc^2(x)).$$

(回想一下, $\cot(x)$ 关于 x 的导数是 $-\csc^2(x)$.) 现在再来看上述原始方程的右边. 对于第一项 $3\csc(y)$, 我们要使用链式求导法则. 我们称该项为 u, 故 $u = 3\csc(y)$. 可以看到 $\mathrm{d}u/\mathrm{d}y = -3\csc(y)\cot(y)$, 因此根据链式求导法则, 有

$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x} = -3\csc(y)\cot(y)\frac{\mathrm{d}y}{\mathrm{d}x}.$$

最后, 项 x^7 关于 x 的导数就是 $7x^6$. 综上, 当对原始方程

$$y\cot(x) = 3\csc(y) + x^7$$

的两边同时关于 x 求导时, 会得到

$$\cot(x)\frac{\mathrm{d}y}{\mathrm{d}x} - y\csc^2(x) = -3\csc(y)\cot(y)\frac{\mathrm{d}y}{\mathrm{d}x} + 7x^6.$$

将所有包含 $\mathrm{d}y/\mathrm{d}x$ 的部分移到等号左边, 而其他的移到等号右边：

$$\cot(x)\frac{\mathrm{d}y}{\mathrm{d}x} + 3\csc(y)\cot(y)\frac{\mathrm{d}y}{\mathrm{d}x} = y\csc^2(x) + 7x^6.$$

现在对等号左边的表达式提取公因式并做除法来求出 $\mathrm{d}y/\mathrm{d}x$：

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y\csc^2(x) + 7x^6}{\cot(x) + 3\csc(y)\cot(y)},$$

这样, 就完成了求解.



最后, 考虑方程

$$x - y\cos\left(\frac{y}{x^4}\right) = \pi + 1.$$

该曲线上点 $(1, \pi)$ 处的切线方程是什么呢？我留给你来做, 代入 $x = 1$ 和 $y = \pi$, 以确信等号两边相等, 从而该点确实在曲线上. 现在, 必须求导. 得到

$$\frac{\mathrm{d}}{\mathrm{d}x}(x) - \frac{\mathrm{d}}{\mathrm{d}x}\left(y\cos\left(\frac{y}{x^4}\right)\right) = \frac{\mathrm{d}}{\mathrm{d}x}(\pi + 1).$$

第一项很容易：它就是 1. 此外, 由于 $\pi + 1$ 是常数, 故等号右边为 0. 剩下的就是中间的一团乱麻. 假设

$$s = y\cos\left(\frac{y}{x^4}\right).$$

那么 s 是 y 和 v 的乘积, 其中 $v = \cos\left(y/x^4\right)$. 根据乘积法则, 有

$$\frac{\mathrm{d}s}{\mathrm{d}x} = v\frac{\mathrm{d}y}{\mathrm{d}x} + y\frac{\mathrm{d}v}{\mathrm{d}x}.$$

真是避无可避：必须要对 v 求导. 我们设 $t = y/x^4$. 那么 $v = \cos(t)$, 故 $\mathrm{d}v/\mathrm{d}t = -\sin(t)$, 并且链式求导法则告诉我们

$$\frac{\mathrm{d}v}{\mathrm{d}x} = \frac{\mathrm{d}v}{\mathrm{d}t}\cdot\frac{\mathrm{d}t}{\mathrm{d}x} = -\sin(t)\frac{\mathrm{d}t}{\mathrm{d}x} = -\sin\left(\frac{y}{x^4}\right)\frac{\mathrm{d}t}{\mathrm{d}x}.$$

不过我们还没有渡过难关, 还需要求出 $\mathrm{d}t/\mathrm{d}x$. 由于 $t = y/x^4$, 我们设 $U = y$ 及 $V = x^4$. (我已经使用了小写的 v, 因此在这里我将使用大写字母.) 商法则告诉我们

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{V\dfrac{\mathrm{d}U}{\mathrm{d}x} - U\dfrac{\mathrm{d}V}{\mathrm{d}x}}{V^2} = \frac{x^4\dfrac{\mathrm{d}y}{\mathrm{d}x} - y\dfrac{\mathrm{d}}{\mathrm{d}x}(x^4)}{(x^4)^2} = \frac{x^4\dfrac{\mathrm{d}y}{\mathrm{d}x} - 4x^3y}{x^8} = \frac{x\dfrac{\mathrm{d}y}{\mathrm{d}x} - 4y}{x^5}.$$

现在只需重新梳理一番. 从后往前推, 可以完成 $\mathrm{d}v/\mathrm{d}x$ 的计算：

$$\frac{\mathrm{d}v}{\mathrm{d}x} = -\sin\left(\frac{y}{x^4}\right)\frac{\mathrm{d}t}{\mathrm{d}x} = -\sin\left(\frac{y}{x^4}\right)\times\frac{x\dfrac{\mathrm{d}y}{\mathrm{d}x} - 4y}{x^5}.$$

这进而能求解 $\mathrm{d}s/\mathrm{d}x$：

$$\frac{\mathrm{d}s}{\mathrm{d}x} = v\frac{\mathrm{d}y}{\mathrm{d}x} + y\frac{\mathrm{d}v}{\mathrm{d}x} = \cos\left(\frac{y}{x^4}\right)\frac{\mathrm{d}y}{\mathrm{d}x} - y\sin\left(\frac{y}{x^4}\right)\times\frac{x\dfrac{\mathrm{d}y}{\mathrm{d}x} - 4y}{x^5}.$$

最后, 回到原始的求导后的方程

$$\frac{\mathrm{d}}{\mathrm{d}x}(x) - \frac{\mathrm{d}}{\mathrm{d}x}\left(y\cos\left(\frac{y}{x^4}\right)\right) = \frac{\mathrm{d}}{\mathrm{d}x}(\pi + 1)$$

它于是可以简化为

$$1 - \cos\left(\frac{y}{x^4}\right)\frac{\mathrm{d}y}{\mathrm{d}x} + y\sin\left(\frac{y}{x^4}\right)\times\frac{x\dfrac{\mathrm{d}y}{\mathrm{d}x} - 4y}{x^5} = 0.$$

没有必要求出 $\mathrm{d}y/\mathrm{d}x$! 我们只想知道当 $x = 1$ 和 $y = \pi$ 时会发生什么. 因此将它们代入. 注意到 $\cos(\pi) = -1$ 及 $\sin(\pi) = 0$, 你应该能看出整个式子可简化为

$$1 - (-1)\frac{\mathrm{d}y}{\mathrm{d}x} + \pi\times 0\times \text{不相关的垃圾} = 0,$$

或 $\mathrm{d}y/\mathrm{d}x = -1$. 因此, 切线方程的斜率是 -1, 并且通过点 $(1, \pi)$, 故其方程为 $y - \pi = -(x - 1)$; 或者如果你喜欢, 可以将它重写成 $y = -x + \pi + 1$.

我们还需要看一下如何对隐函数求二阶导的问题. 不过在此之前, 让我们小结一下前面所用的方法：

- 在原始方程中, 对一切求导并使用链式求导法则、乘积法则以及商法则进行化简;
- 如果想要求 $\mathrm{d}y/\mathrm{d}x$, 可重新整理并作除法来求解 $\mathrm{d}y/\mathrm{d}x$; 不过

- 如果想要求的是斜率或求曲线一个特定点上的切线方程, 可先代入 x 和 y 的已知值, 接着重新整理并求 $\mathrm{d}y/\mathrm{d}x$, 然后如果需要的话, 使用点斜式来求切线方程.

8.1.2 隐函数求二阶导



求导两次可以得到二阶导. 例如, 如果

$$2y+\sin(y)=\frac{x^2}{\pi}+1,$$

那么该曲线上点 $(\pi,\pi/2)$ 处的 $\mathrm{d}^2y/\mathrm{d}x^2$ 的值是什么呢? 再一次地, 你应该先通过代入 x 和 y 的值, 看看方程是否成立来检验该点是否位于曲线上. 现在, 如果你想要求导两次, 必须先从求导一次开始! 使用链式求导法则来处理 $\sin(y)$ 这一项, 你应该会得到

$$2\frac{\mathrm{d}y}{\mathrm{d}x}+\cos(y)\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{2x}{\pi},$$

现在, 需要再求导一次. 务必先不要做代入! 为了求导, 需要查看当 x 和 y 变化时会有什么情况发生. 如果固定了它们的值 (如 π 和 $\pi/2$), 就不可能看到变化情况了. 相反, 对上述方程关于 x 求导:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(2\frac{\mathrm{d}y}{\mathrm{d}x}\right)+\frac{\mathrm{d}}{\mathrm{d}x}\left(\cos(y)\frac{\mathrm{d}y}{\mathrm{d}x}\right)=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{2x}{\pi}\right).$$

等号右边正好是 $2/\pi$, 左边第一项正好是 $2\left(\mathrm{d}^2y/\mathrm{d}x^2\right)$. 棘手的是左边第二项. 我们需要使用乘积法则: 设 $s=\cos(y)(\mathrm{d}y/\mathrm{d}x)$, 以及 $u=\cos(y)$ 和 $v=\mathrm{d}y/\mathrm{d}x$, 这样 $s=uv$. 根据乘积法则,

$$\frac{\mathrm{d}s}{\mathrm{d}x}=v\frac{\mathrm{d}u}{\mathrm{d}x}+u\frac{\mathrm{d}v}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}x}\cdot\frac{\mathrm{d}u}{\mathrm{d}x}+\cos(y)\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)=\frac{\mathrm{d}y}{\mathrm{d}x}\cdot\frac{\mathrm{d}u}{\mathrm{d}x}+\cos(y)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}.$$

我们仍需要求出 $\mathrm{d}u/\mathrm{d}x$, 其中 $u=\cos(y)$. 这其实就是链式求导法则的再次运用:

$$\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{\mathrm{d}u}{\mathrm{d}y}\cdot\frac{\mathrm{d}y}{\mathrm{d}x}=-\sin(y)\frac{\mathrm{d}y}{\mathrm{d}x}.$$

综上, 可以看到

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}x}&=\frac{\mathrm{d}y}{\mathrm{d}x}\cdot\frac{\mathrm{d}u}{\mathrm{d}x}+\cos(y)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}y}{\mathrm{d}x}\cdot\left(-\sin(y)\frac{\mathrm{d}y}{\mathrm{d}x}\right)+\cos(y)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\\&=-\sin(y)\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2+\cos(y)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}.\end{aligned}$$

注意: 量

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 \quad 和 \quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2}$$

是完全不同的! 左边的量是一阶导的平方, 而右边的量是二阶导. 不管怎样, 让我们把一切凑在一起. 先从

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(2\frac{\mathrm{d}y}{\mathrm{d}x}\right)+\frac{\mathrm{d}}{\mathrm{d}x}\left(\cos(y)\frac{\mathrm{d}y}{\mathrm{d}x}\right)=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{2x}{\pi}\right)$$

开始, 现在可以将它写成

$$2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}-\sin(y)\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2+\cos(y)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{2}{\pi}.$$

呼, 可费老劲了. 不过这还没算完：仍需要求出当 $x=\pi$ 和 $y=\pi/2$ 时的 $\mathrm{d}^2y/\mathrm{d}x^2$. 因此, 将它们代入上述方程, 会得到

$$2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}-\sin\left(\frac{\pi}{2}\right)\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2+\cos\left(\frac{\pi}{2}\right)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{2}{\pi}.$$

该式简化为

$$2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}-\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2=\frac{2}{\pi}.$$

问题是, 我们仍需要知道 $\mathrm{d}y/\mathrm{d}x$! 但这不成问题：在方程

$$2\frac{\mathrm{d}y}{\mathrm{d}x}+\cos(y)\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{2x}{\pi}$$

中代入 $x=\pi$ 和 $y=\pi/2$, (我之前一直没有让你这么做!) 会得到

$$2\frac{\mathrm{d}y}{\mathrm{d}x}+0\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{2\pi}{\pi}=2,$$

因此, $\mathrm{d}y/\mathrm{d}x=1$. 将其代入二阶导方程, 得到

$$2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}-(1)^2=\frac{2}{\pi}.$$

这意味着, 当 $x=\pi$ 且 $y=\pi/2$ 时,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{1}{\pi}+\frac{1}{2}.$$

由此, 我们终于完成了求解!

8.2　相关变化率

设想有两个相关的量 (随你喜欢让它们测量什么), 如果你知道其中之一, 就可以求出另外一个. 例如, 如果你一直盯着一架飞过你头顶的飞机, 那么你的视线与地面的夹角取决于飞机的位置. 在这种情况下, 这两个量就是飞机的位置和我刚刚描述的那个夹角.

当然, 当这两个量中的一个发生变化时, 另一个也会发生相应的变化. 假设我们知道其中一个量变化有多快, 那么另一个量的变化有多快呢？这就是我们所说**相**

关变化率的意思. 你看, **变化率**是一个量随时间改变的速率. 现在我们有两个相关的量, 想要知道它们的变化率是如何相互关联的.

以上变化率的定义有点粗略. 如果你想要知道某物随时间的变化有多快, 只需简单地对其关于时间求导. 因此, 以下是其真正的定义: **量 Q 的变化率是 Q 关于时间的导数**. 也就是说,

$$\boxed{\text{如果 } Q \text{ 是某个量, 那么 } Q \text{ 的变化率是 } \frac{\mathrm{d}Q}{\mathrm{d}t}.}$$

当你看到 "变化率" 这几个字时, 应该自动想到 "$\mathrm{d}/\mathrm{d}t$".

那么如何从一个涉及两个相关量的方程求出涉及这两个量的相关变化率的方程呢? 当然是求导了! 如果你对等号两边关于 t 做隐函数求导的话, 就会发现相关变化率跃然而出, 给出一个新的方程. 如果你面对三个或更多的相关量 (例如, 一个矩形的长度、宽度和面积), 这样做同样有效. 对其关于 t 做隐函数求导, 就会将各个变化率关联在一起.

先让我们看看求解相关变化率问题的一般方法. 然后, 我们会用它来求解一大堆的例子.

(1) 读题. 识别出所有的量并注意到哪一个量是你需要对其求相关变化率的. 如果需要的话, 可以画图!

(2) 写出一个关联所有量的方程 (有时候你需要不止一个方程). 为了做到这一步, 你可能需要用到几何学, 可能是涉及相似三角形的. 如果你有不止一个方程, 试着把它们联立求解, 以消去不必要的变量.

(3) 对剩余的方程关于时间 t 做隐函数求导. 也就是说, 每一个方程两边各添加一个 $\dfrac{\mathrm{d}}{\mathrm{d}t}$. 你会得到一个或多个关联起各个变化率的方程.

(4) 最后, 将你所知道的值代入所有的方程中做替换. 联立求解方程得到你想要的变化率.

这类问题与你以前所见的应用题的唯一区别在于第 3 步. 在这里, 它让一切完全不同. 在看例子之前, 还要提醒一点: **最后才做值的替换, 要在求导之后!** 这就是说, 不要调换第 3 步和第 4 步. 如果你先做替换, 就会让量无从变化, 从而变化率将全部为 0. 冻结一切, 你就只能得到这个了 ······

8.2.1 一个简单的例子

下面是一个说明上述方法的相对简单的例子. 设想用打气筒给一个完美球体的气球充气. 空气以常数速率 12π 立方英寸每秒进入气球. 当气球的半径达到 2 英寸时, 气球的半径的变化率是多少? 此外, 当气球的体积达到 36π 立方英寸时, 气球的半径的变化率又是多少?

好, 先让我们写出全部的量 (第 1 步). 它们分别是气球的体积和半径. 我们称体积为 V(单位是立方英寸), 半径为 r(单位是英寸). 我们需要求出半径 r 的相关变化率. 现在, 需要一个关联 V 和 r 的方程 (第 2 步). 这里会用到一些几何知识. 由于气球是一个球体, 我们知道

$$V = \frac{4}{3}\pi r^3.$$

该式关联了所有的量. 现在, 需要关联各个变化率了 (第 3 步). 对方程两边关于 t 做隐函数求导:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{4}{3}\pi r^3\right).$$

左边正好是 $\mathrm{d}V/\mathrm{d}t$; 为了处理右边, 令 $s = r^3$, 这样 $\mathrm{d}s/\mathrm{d}r = 3r^2$. 根据链式求导法则,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\mathrm{d}s}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}t} = 3r^2\frac{\mathrm{d}r}{\mathrm{d}t}.$$

现在, 可以将它代入上述方程, 得到

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{4}{3}\pi\left(3r^2\frac{\mathrm{d}r}{\mathrm{d}t}\right) = 4\pi r^2\frac{\mathrm{d}r}{\mathrm{d}t}.$$

这样, 就有了一个关联 V 和 r 的变化率的方程. 最后准备做替换 (第 4 步). 在问题的两个部分中, 体积的变化率都是 12π 立方英寸/秒. 用符号表示, 我们有 $\mathrm{d}V/\mathrm{d}t = 12\pi$. 将它代入上述方程, 得到

$$12\pi = 4\pi r^2\frac{\mathrm{d}r}{\mathrm{d}t}.$$

整理可得

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{3}{r^2}.$$

太棒了! 这意味着, 如果我们知道半径 r, 那么就可以求出半径的变化率, 也就是 $\mathrm{d}r/\mathrm{d}t$. 注意到半径的变化率本身是一个变化的量, 它依赖于半径. 你很可能也经历过, 当你吹气球时, 一开始它的大小 (或半径) 会增长得很快, 然后其增长速度会降低, 尽管你一直是将相同量的空气吹进气球的. 上述 $\mathrm{d}r/\mathrm{d}t$ 的公式验证了这一点, 它在 r 上是递减的.

有了这个公式, 我们可以快速地对问题的两个部分作解答. 对于第一部分, 我们知道半径是 2 英寸, 故在以上公式中令 $r = 2$, 得到

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{3}{2^2} = \frac{3}{4}.$$

因此答案为 $\dfrac{3}{4}$. 但 $\dfrac{3}{4}$ 什么呢? 所以有必要写一句话总结一下, 也别忘了测量的单位. 在这种情况下, 我们会说, 当半径达到 2 英寸时, 半径的变化率是 $\dfrac{3}{4}$ 英寸/秒.

对于问题的第二部分, 我们知道体积是 36π 立方英寸. 这意味着 $V = 36\pi$. 问题是, 为了求出 $\mathrm{d}r/\mathrm{d}t$, 我们需要知道 r 是什么. 现在, 我们需要回到关联 V 和 r 的方程, 也是 $V = \frac{4}{3}\pi r^3$. 如果将 $V = 36\pi$ 代入并求解 r, 应该可以看出 $r = 3$ 英寸. 最后, 将它代入到 $\mathrm{d}r/\mathrm{d}t$ 的方程中, 得出

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{3}{r^2} = \frac{3}{3^2} = \frac{1}{3}.$$

因此, 当体积达到 36π 立方英寸时, 半径的变化率是 $\frac{1}{3}$ 英寸/秒.

8.2.2 一个稍难的例子

让我们来看看另一个还相对简单的例子, 这一次涉及三个量. 假设有两辆汽车 A 和 B. 汽车 A 在一条路上径直向北行驶远离你家, 而汽车 B 在另一条路上径直向西行驶接近你家. 汽车 A 以 55 英里/小时的速度行驶, 而汽车 B 以 45 英里/小时的速度行驶. 当 A 到达你家北面 21 英里, 而 B 到达你家东面 28 英里时, 两辆汽车间的距离的变化率是多少?

为了回答这个问题, 我们最好来画图 (第 1 步). 画出你家 H 以及汽车 A 和 B. 令 H 和 A 间的距离为 a, H 和 B 间的距离为 b, 而令两辆汽车间的距离为 c, 如图 8-1 所示.

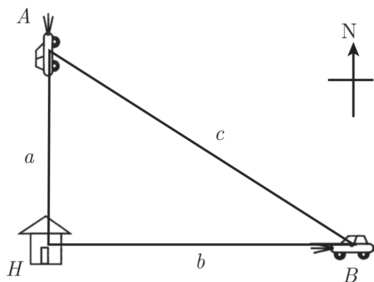


图 8-1

注意到不好用 21 替代 a, 或用 28 替代 b. 你想看的是当 a 和 b 变化时会发生什么, 而不是当它们固定在某个特定的数时会发生什么, 因此它们需要有作为变量的可变性. 还要注意到 c 是我们想要对其求变化率的量, 因为它就是两辆汽车间的距离.

接下去是第 2 步. 关联 a、b 和 c 的方程不是别的, 正是勾股定理:

$$a^2 + b^2 = c^2.$$

进入第 3 步, 对其关于 t 做隐函数求导, 得到

$$2a\frac{\mathrm{d}a}{\mathrm{d}t} + 2b\frac{\mathrm{d}b}{\mathrm{d}t} = 2c\frac{\mathrm{d}c}{\mathrm{d}t}.$$

我们知道, 汽车 A 正以 55 英里/小时的速度远离你家. 这意味着, 距离 a 是以 55 英里/小时的速度而增加的, 因此 $\mathrm{d}a/\mathrm{d}t = 55$. 至于 B, 它正以 45 英里/小时的速度接近你家. 这意味着, 距离 b 是以 45 英里/小时的速度而减少的, 因此 $\mathrm{d}b/\mathrm{d}t = -45$. 这里你需要一个负号! 否则, 你会搞砸整个求解过程. 将这些值代入上述方程, 我们得到

$$2a(55) + 2b(-45) = 2c\frac{\mathrm{d}c}{\mathrm{d}t},$$

它可以被简化为

$$c\frac{\mathrm{d}c}{\mathrm{d}t} = 55a - 45b.$$

最后, 可以看到我们感兴趣的时刻即当 $a = 21$ 和 $b = 28$ 时所发生的情况. 在那一时刻, 我们知道 $c^2 = 21^2 + 28^2$, 即 $c = \pm 35$. 由于 c 是正的, (它是两辆汽车间的距离!) 因而有 $c = 35$. 将那些数代入上述方程, 得到

$$(35)\frac{\mathrm{d}c}{\mathrm{d}t} = 55(21) - 45(28).$$

通过从等式两边消去因子 5 和 7, 可以很容易地进行计算. 最后的结果是 $\mathrm{d}c/\mathrm{d}t = -3$. 这意味着, 在我们所考虑的那一时刻, 两辆汽车间的距离是以 3 英里/小时的变化率**减少**的.

这就是我们需要的答案了. 注意到在我们所考虑的那一时刻, 两辆汽车实际上越来越接近, 尽管 A 以比 B 更快的速度远离你家. 如果我们等待一小会儿, 汽车 A 会离你家更远, 而汽车 B 会更加接近你家; 由 $\mathrm{d}c/\mathrm{d}t$ 的方程可知, 这个量终究会变成正的 (尽管问题没有涉及这一点).

8.2.3　一个更难的例子



下面是一个更难的涉及相似三角形的例子：设想有一个奇怪的巨大的圆锥形水罐 (锥尖在下方). 圆锥的高是圆锥半径的两倍. 如果水是以 8π 立方英尺/秒的速率注入水罐, 求当水罐中水的体积为 18π 立方英尺时, 水位的变化率是多少?

这个问题也有第二部分：设想水罐底部有一个小洞, 致使水罐中每一立方英尺的水以一立方英尺每秒的速率流出. 我想知道同样的事情：当水罐中的水的体积为 18π 立方英尺时, 水位的变化率是多少 (但现在水罐有洞)?

让我们从第一部分开始. 情形如图 8-2 所示.

我们在图中标记了一些量. 水罐的高为 H, 其半径为 R. 水位的高度为 h, 水位顶部水平面的半径为 r. 所有这些量的测量单位都是英尺. 我们还令 v 是水罐中水的体积, 测量单位是立方英尺. (你可以令 V 是整个水罐的体积, 但我们从来不需要这个量, 因为水罐决不会被灌满 —— 它就是那么大!) 不管怎样, 这完成了第 1 步.

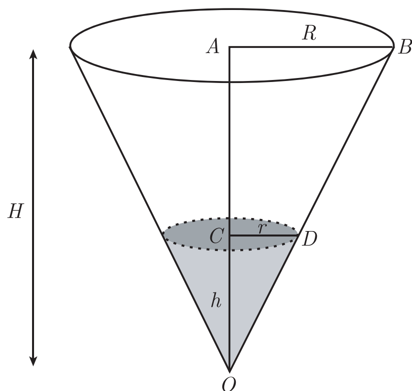


图 8-2

对于第 2 步, 必须开始关联那些量中的某些量了. 我们已知水罐的高是半径的两倍, 因此有 $H = 2R$. 尽管如此, 我们对关联 h 和 r 更感兴趣. 在图中有一些相似三角形：事实上, $\triangle ABO$ 和 $\triangle CDO$ 相似, 故 $H/R = h/r$. 由于 $H = 2R$, 因而 $2R/R = h/r$, 就是说 $h = 2r$. 因此, 水罐中的水就像是整个水罐的微缩复制. 不管怎样, 我们仍需要求出用 h 和 r 表示的水罐中水的体积. 高为 h 单位、半径为 r 单位的圆锥的体积由公式 $v = \dfrac{1}{3}\pi r^2 h$ 立方单位给出. 在这里, 消去 h 和 r 中的一个会很好, 又由于我们对水位 h 比半径 r 更感兴趣, (通过读题会知道为什么!) 所以消去 r 会更有意义. 代入 $r = h/2$, 有

$$v = \frac{1}{3}\pi r^2 h = \frac{1}{3}\pi\left(\frac{h}{2}\right)^2 h = \frac{\pi h^3}{12}, \quad \text{即} \quad v = \frac{\pi h^3}{12}.$$

现在, 对于第 3 步, 对上式关于 t 求导. 根据链式求导法则,

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\pi}{12} \times 3h^2 \frac{\mathrm{d}h}{\mathrm{d}t} = \frac{\pi h^2}{4}\frac{\mathrm{d}h}{\mathrm{d}t}, \quad \text{即} \quad \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\pi h^2}{4}\frac{\mathrm{d}h}{\mathrm{d}t}.$$

很好! 现在来看第 4 步, 将我们所知的一切代入以上两个方程中. 我们知道 $\mathrm{d}v/\mathrm{d}t = 8\pi$, 并且对当 $v = 18\pi$ 时会发生什么感兴趣. 分别作替换, 得到

$$18\pi = \frac{\pi h^3}{12} \quad \text{和} \quad 8\pi = \frac{\pi h^2}{4}\frac{\mathrm{d}h}{\mathrm{d}t}.$$

第一个方程告诉我们 $h^3 = 18 \times 12 = 216$, 故 $h = 6$. 也就是说, 当水的体积达到 18π 立方英尺时, 水位是 6 英尺. 将其代入第二个方程, 得到

$$8\pi = \frac{\pi}{4} \times 6^2 \frac{\mathrm{d}h}{\mathrm{d}t},$$

这意味着 $\mathrm{d}h/\mathrm{d}t = 8/9$. 也就是说, 在我们关心的时刻 (当水的体积达到 18π 立方英尺时), 水位以 8/9 英尺/秒的速率上升.

第二部分几乎是一样的. 事实上, 唯一的区别出现在第 4 步. 我们仍然想用

$v = 18\pi$ 作替换, 这将意味着再次有 $h = 6$. 另一方面, 代入 $\mathrm{d}v/\mathrm{d}t = 8\pi$ 是错误的, 因为这根本没有考虑到那个洞. 我们知道每秒有 8π 立方英尺的水注入罐中, 但对于罐中每一立方英尺的水来说, 每秒有一立方英尺的水流出来. 由于在罐中有 v 立方英尺的水, (由定义可知!) 从洞中流出的水的速率是 v 立方英尺每秒. 因此, 流入水的速率是 8π, 而流出水的速率是 v (它们的单位都是立方英尺每秒), 这意味着

$$\frac{\mathrm{d}v}{\mathrm{d}t} = 8\pi - v.$$

现在, 当 $v = 18\pi$ 时, 我们有 $\mathrm{d}v/\mathrm{d}t = 8\pi - 18\pi = -10\pi$. 因此, 需要将 $\mathrm{d}v/\mathrm{d}t = -10\pi$ 和 $h = 6$ 代入先前的方程

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\pi h^2}{4}\frac{\mathrm{d}h}{\mathrm{d}t},$$

结果是 $\mathrm{d}h/\mathrm{d}t = -10/9$. 这意味着, 在我们所考虑的那一时刻, 罐中的水位以 10/9 英尺每秒的速率下降. 尽管我们正在向水罐注水, 但洞会让更多的水流出并导致水位下降.

8.2.4　一个非常难的例子

这里还有一个问题. 你已经看过不少相关变化率的问题, 现在或许应该尝试一下在读答案之前自己求解.

设想有一架飞机保持在 2000 英尺的高度远离你朝正东方向飞行. 飞机以 500 英尺每秒的常数速率飞行. 同时, 不久之前有一个跳伞员从直升飞机 (它已经飞走了) 上跳下来. 跳伞员在你东边 1000 英尺处上空垂直地以 10 英尺每秒的常数速率向下飘落. 情形如图 8-3 所示.

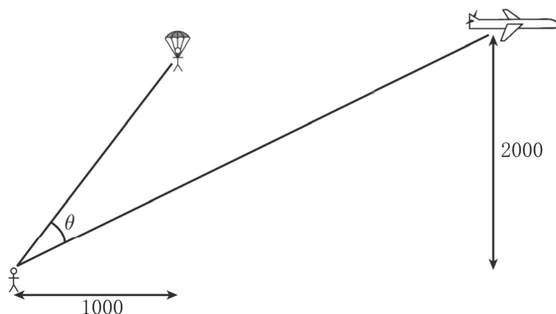


图　8-3

在图中, 跳伞员相对于你的方位角与飞机相对于你的方位角之差被标记为 θ. 问题是, 当飞机和跳伞员在同一高度, 但飞机在你东边 8000 英尺时, 角 θ 的变化率是多少?

我们有两个需要关心的对象, 即飞机和跳伞员. 我们知道, 飞机的飞行高度总是 2000 英尺 (相对于你的脑袋), 但我们不知道飞机在你东边多远 —— 距离总是

在变化的. 令飞机在你东边 p 英尺. 至于跳伞员, 这一次我们确切地知道跳伞员在你东边到底多远 ——1000 英尺. 但问题是, 跳伞员的高度是多少呢? 令其高度为 h 英尺. 通过画几条辅助线, 可以将图改写成图 8-4.

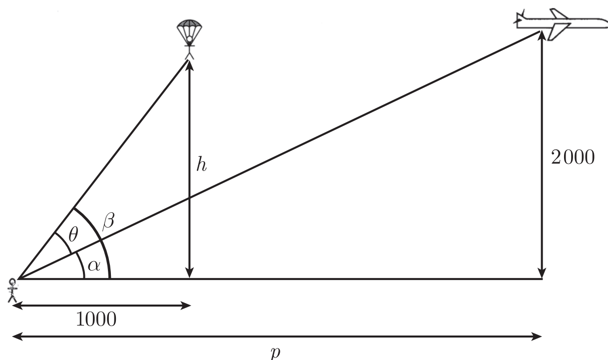


图 8-4

注意到量 1000 和 2000 绝不会变化, 但量 p 和 h 会变化. 特别是, 飞机向右飞行, 因此 p 会越来越大; 跳伞员向下移动, 因此 h 会越来越小. 尽管问题让我们关注 $p = 8000$ 和 $h = 2000$ (和飞机有相同的高度) 的那一时刻, 但我们必须允许 p 和 h 发生变化, 以便可以算出变化率. 毕竟, 如果 p 和 h 保持不变, 那么飞机和跳伞员就悬停在空中的同一个地方, 当然角 θ 也不会改变. 这是很不现实的 —— 因此需要让 p 和 h 变化, 从而角 θ 也会变化, 而我们可以算出它变化得有多快. 这就完成了第 1 步.

说到角 θ, 从图中很明显, 它就是跳伞员和地面的夹角 β 与飞机和地面的夹角 α 之差. (我们假设你没有高度, 或者换个说法, 你是躺在地面上的.) 因此, 我们知道 $\theta = \beta - \alpha$. 事实上, 应该写成 $\theta = |\beta - \alpha|$, 以防跳伞员低于飞机. 不过在我们感兴趣的时刻附近, 高度是一样的, 而飞机比跳伞员还要偏东, 因此 β 一定大于 α, 我们不需要绝对值.

现在, 要做一些三角函数运算. 我们有两个直角三角形. 从它们中的一个 (有飞机的那个), 得到 $\tan(\alpha) = 2000/p$. 从另一个, 得到 $\tan(\beta) = h/1000$. 我们将这些方程写在一起:

$$\tan(\alpha) = \frac{2000}{p} \quad \text{和} \quad \tan(\beta) = \frac{h}{1000}.$$

第 2 步终于结束了, 现在可以进入第 3 步, 对这两个关系关于时间做隐函数求导. 由第一个开始, 令 $u = \tan(\alpha)$, $v = 2000/p$, 这样方程就变为 $u = v$. 这意味着, $\mathrm{d}u/\mathrm{d}t = \mathrm{d}v/\mathrm{d}t$. 我们使用链式求导法则来求这些量. 首先是 $\mathrm{d}u/\mathrm{d}t$:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\mathrm{d}u}{\mathrm{d}\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \sec^2(\alpha)\frac{\mathrm{d}\alpha}{\mathrm{d}t}.$$

接着是 $\mathrm{d}v/\mathrm{d}t$：

$$\frac{\mathrm{d}v}{\mathrm{d}t}=\frac{\mathrm{d}v}{\mathrm{d}p}\frac{\mathrm{d}p}{\mathrm{d}t}=-\frac{2000}{p^2}\frac{\mathrm{d}p}{\mathrm{d}t}.$$

由于 $\mathrm{d}u/\mathrm{d}t=\mathrm{d}v/\mathrm{d}t$, 有

$$\sec^2(\alpha)\frac{\mathrm{d}\alpha}{\mathrm{d}t}=-\frac{2000}{p^2}\frac{\mathrm{d}p}{\mathrm{d}t}.$$

这只是两个三角方程的第一个. 对于第二个涉及 $\tan(\beta)$ 的方程, 我们需要重复刚才的做法. 方程左边与处理 $\tan(\alpha)$ 时是一样的, 但右边要更简单些. 你应该弄明白我们如何得到

$$\sec^2(\beta)\frac{\mathrm{d}\beta}{\mathrm{d}t}=\frac{1}{1000}\frac{\mathrm{d}h}{\mathrm{d}t}.$$

回想一下, 我们还知道 $\theta=\beta-\alpha$, 因此也可以对其关于时间 t 求导, 并得到 $\mathrm{d}\theta/\mathrm{d}t=\mathrm{d}\beta/\mathrm{d}t-\mathrm{d}\alpha/\mathrm{d}t$. 方程太多了, 让我们将所有的六个写在一起：

$$\begin{aligned}
\tan(\alpha)&=\frac{2000}{p} &\qquad \sec^2(\alpha)\frac{\mathrm{d}\alpha}{\mathrm{d}t}&=-\frac{2000}{p^2}\frac{\mathrm{d}p}{\mathrm{d}t}\\
\tan(\beta)&=\frac{h}{1000} &\qquad \sec^2(\beta)\frac{\mathrm{d}\beta}{\mathrm{d}t}&=\frac{1}{1000}\frac{\mathrm{d}h}{\mathrm{d}t}\\
\theta&=\beta-\alpha &\qquad \frac{\mathrm{d}\theta}{\mathrm{d}t}&=\frac{\mathrm{d}\beta}{\mathrm{d}t}-\frac{\mathrm{d}\alpha}{\mathrm{d}t}.
\end{aligned}$$

现在, 最好做一些替换, 从这一团混乱中找出真相. 我们知道什么呢? 飞机的速率是 500 英尺每秒, 这意味着, $\mathrm{d}p/\mathrm{d}t=500$. 跳伞员的速率是 10 英尺每秒, 但其高度是递减的, 故 $\mathrm{d}h/\mathrm{d}t=-10$. 如果你忘记了这个负号, 就会得到错误的答案! 因此要特别小心. 例如, 如果飞机是朝向你飞行的, 那么 p 将是递减的, 故 $\mathrm{d}p/\mathrm{d}t$ 将是负的. 不管怎样, 我们感兴趣的时刻是飞机在 8000 英尺以外, 故 $p=8000$, 以及跳伞员的高度为 2000 英尺, 故 $h=2000$. 前四个方程变得简单多了：

$$\begin{aligned}
\tan(\alpha)&=\frac{2000}{8000}=\frac{1}{4} &\qquad \sec^2(\alpha)\frac{\mathrm{d}\alpha}{\mathrm{d}t}&=-\frac{2000}{8000^2}\times 500=-\frac{1}{64}\\
\tan(\beta)&=\frac{2000}{1000}=2 &\qquad \sec^2(\beta)\frac{\mathrm{d}\beta}{\mathrm{d}t}&=\frac{1}{1000}\times(-10)=-\frac{1}{100}.
\end{aligned}$$

只要再知道 $\sec^2(\alpha)$ 是何值, 就可以由右上角的方程求出 $\mathrm{d}\alpha/\mathrm{d}t$. 但等一下, 我们已经知道 $\tan(\alpha)=1/4$, 显然可以求出 $\sec^2(\alpha)$. 根据三角恒等式 (参见 2.4 节), 我们得到

$$\sec^2(\alpha)=1+\tan^2(\alpha)=1+\left(\frac{1}{4}\right)^2=\frac{17}{16}.$$

因此, 右上角的方程变为

$$\frac{17}{16}\frac{\mathrm{d}\alpha}{\mathrm{d}t} = -\frac{1}{64},$$

结果为

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = -\frac{1}{68}.$$

这太棒了！现在需要再对 β 做相同的操作, 然后就会完成求解. 这里, 我们知道 $\tan(\beta) = 2$, 故

$$\sec^2(\beta) = 1 + \tan^2(\beta) = 1 + 2^2 = 5.$$

将其代入上述右下角的方程中, 得到

$$5\frac{\mathrm{d}\beta}{\mathrm{d}t} = -\frac{1}{100},$$

这意味着

$$\frac{\mathrm{d}\beta}{\mathrm{d}t} = -\frac{1}{500}.$$

因此我们知道了 $\mathrm{d}\alpha/\mathrm{d}t$ 和 $\mathrm{d}\beta/\mathrm{d}t$ 的值, 进而从前述原始六个方程中的最后一个可知,

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\mathrm{d}\beta}{\mathrm{d}t} - \frac{\mathrm{d}\alpha}{\mathrm{d}t} = \left(-\frac{1}{500}\right) - \left(-\frac{1}{68}\right) = \frac{-17+125}{8500} = \frac{27}{2125}.$$

因此在我们所考虑的那一时刻, 角 θ 是以 $27/2125$ 弧度每秒的速率递增的. 我们终于完成了求解.

第 9 章　指数函数和对数函数

这巨长的一章都是关于指数函数和对数函数的. 在回顾完这些函数的性质之后, 我们需要对它们做一些微积分的运算. 事实表明, 有一个特殊的底数, 数 e, 它的相关性质相当好. 特别是, 对 e^x 和 $\log_{\mathrm{e}}(x)$ 做微积分的运算要比处理像 2^x 和 $\log_3(x)$ 这样的量稍微简单些. 因此, 我们需要花一些时间来看看 e. 还有其他一些情况我们也想看看; 总之, 本章计划讨论下列话题：

- 回顾指数函数和对数函数的基本知识, 以及两者是如何相互关联的;
- e 的定义和性质;
- 如何对指数函数和对数函数求导;
- 如何求解涉及指数函数和对数函数的极限问题;
- 对数函数的微分;
- 指数增长和指数衰变;
- 双曲函数.

9.1　基 础 知 识

在开始对指数函数和对数函数做微积分的运算之前, 你真的需要理解它们的性质. 简单来说, 除了对数函数的真正定义之外, 你还需要知道三点：指数法则、对数和指数的关系, 以及对数法则.

9.1.1　指数函数的回顾

这里的大致思想是, 我们取一个正数, 称之为**底数**, 并将它提升为其**指数**次方的一个幂：

$$\text{底数}^{\text{指数}}.$$

例如, 数 $2^{-5/2}$ 是一个底数为 2、指数为 $-5/2$ 的幂. 重要的是, 你要知道所谓的指数法则, 它们实际上告诉了你指数函数是如何运算的. 毫无疑问, 你已经见过它们了, 但在这里重新列出, 以便再次提醒你. 对于任意的底数 $b>0$ 和实数 x 与 y：

(1) $\boxed{b^0=1.}$ 任意非零数的零次幂是 1.

(2) $\boxed{b^1=b.}$ 一个数的一次幂正好是该数本身.

(3) $\boxed{b^xb^y=b^{x+y}.}$ 当将两个底数相同的幂相乘时, 将指数**相加**.

(4) $\boxed{\dfrac{b^x}{b^y}=b^{x-y}.}$ 当将两个底数相同的幂相除时, 将分子的指数**减去**分母的

指数.

(5) $\boxed{(b^x)^y = b^{xy}.}$ 当取幂的幂时, 将指数**相乘**.

你也应该知道指数函数的图像是什么样子的. 我们已经在 1.6 节粗略见过, 但不管怎样, 我们将很快再次讨论到其图像.

9.1.2 对数函数的回顾

对数 —— 一个让许多学生闻名丧胆的词. 看仔细了, 现在我们就来看看如何直面这些怪兽. 设想, 你想要从方程

$$2^x = 7$$

中求解 x. 将 x 从指数的位置移下来的方法是在方程两边取对数. 由于左边的底数是 2, 对数的底就是 2. 事实上, 根据定义, 上述方程的解就是

$$x = \log_2(7).$$

换句话说, 必须将 2 提升为其几次幂才能得到 7 呢? 答案是 $\log_2(7)$. 这个特定的数不能被简化, 但 $\log_2(8)$ 呢? 问问自己, 必须将 2 提升为其几次幂才能得到 8? 由于 $2^3 = 8$, 我们需要的幂次就是 3. 因此, $\log_2(8) = 3$.

回到方程 $2^x = 7$. 我们已经知道这意味着 $x = \log_2(7)$. 而如果现在将 x 的值代入原始方程中, 将得到下面这个看起来很奇怪的公式:

$$2^{\log_2(7)} = 7.$$

更一般地, **$\log_b(y)$ 是为了得到 y 你必须将底数 b 提升的幂次**. 这意味着, 对于给定的 b 和 y, $x = \log_b(y)$ 是方程 $b^x = y$ 的解. 将 x 的值代入, 得到公式

$$\boxed{b^{\log_b(y)} = y.}$$

它对于任意的 $y > 0$ 和 $b > 0$ (除了 $b = 1$) 都成立. 但为什么我要坚持让 b 和 y 都是正的呢? 首先, 如果 b 是负的, 那么很多怪诞的事情就会发生. 量 b^x 可能就没有定义了. 例如, 如果 $b = -1$ 且 $x = 1/2$, 那么 b^x 就是 $(-1)^{1/2}$, 它是 $\sqrt{-1}$. (真糟糕!) 因此, 为了避免所有这些, 我们要求 $b > 0$. 这样, b 取任意次幂就没有问题了. 另一方面, b^x 总是正的! 因此, 如果 $y = b^x$, 那么一定有 $y > 0$. 这意味着, 取一个负数或 0 的对数是毫无意义的. 毕竟, 如果 $\log_b(y)$ 是为了得到 y 你必须将底数 b 提升的幂次, 那你就不可能将 b 提升为其几次幂而得到一个负数或 0, 于是 y 不可能是负数或 0. **你只能取一个正数的对数**.

你或许也已经注意到, 我提到 $b = 1$ 不好. 如果你将 $b = 1$ 代入上述公式 $b^{\log_b(y)} = y$, 会得到 $1^{\log_1(y)} = y$. 但问题是, 我将 1 提升为其任意次幂的结果仍然是 1, 但 y 可能不是 1, 因此这个方程说不通. 也就是说, 根本就不存在底数为 1 的对数. 那么底数为 1/2 呢? 这没问题, 但我们很少需要一个底数为 1/2 的对数, 因为事实表明, 对于任意的 y, $\log_{1/2}(y) = -\log_2(y)$. (通过设 $y = (1/2)^x$ 并注意到 y

也等于 2^{-x}, 你便可以证明该式.) 同理, 对于任意的介于 0 和 1 的底数 b, 对于所有的 y, $\log_b(y) = -\log_{1/b}(y)$, 且 $1/b$ 大于 1. 因此, 从现在开始, 我们将总是假设底数 b 大于 1.

9.1.3　对数函数、指数函数及反函数

通过使用反函数, 我们可以对之前看到的一切进行更精密的描述. 固定一个底数 $b > 1$ 并且设 $f(x) = b^x$. 函数 f 的定义域是 $\mathbb{R}$ 且值域为 $(0, \infty)$. 由于它通过了水平线检验, 因此它有反函数, 我们称之为 g. g 的定义域是 f 的值域, 即 $(0, \infty)$, 而 g 的值域就是 f 的定义域, 即 $\mathbb{R}$. 我们说, g 是**底数为 b 的对数**. 事实上, 根据定义, $g(x) = \log_b(x)$. 忆及反函数的图像就是原始函数关于镜面直线 $y = x$ 的映像, 我们可以在同一坐标系下画出 $f(x) = b^x$ 及其反函数 $g(x) = \log_b(x)$ 的图像, 如图 9-1 所示.

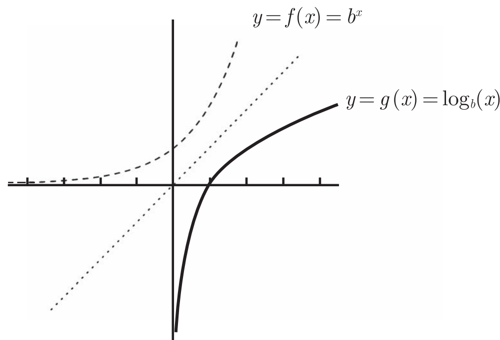


图　9-1

由于 f 和 g 互为反函数, 我们知道 $f(g(x)) = x$, $g(f(x)) = x$. (正如我们将要看到的, 第一个事实仅在 $x > 0$ 时成立.) 让我们对这两个事实一一进行阐释.

(1) 从 $f(g(x)) = x$ 开始. 由于 g 是对数函数, 因此 x 最好是正的 (回想一下, 你只能取一个正数的对数.) 现在, 我们来仔细看看量 $f(g(x))$. 你以一个正数 x 开始, 将它代入 g 中, g 是底数为 b 的对数. 然后, 再对结果进行指数运算, 即将 b 提升为其 $g(x)$ 次幂. 结果会得到原始的数! 事实上, 由于 $f(x) = b^x$ 且 $g(x) = \log_b(x)$, 公式 $f(g(x)) = x$ 其实说的是

$$b^{\log_b(x)} = x,$$

也就是上一节中的其中一个公式 (用 x 替换了 y). 只要底数相同, **对数的指数就是原始的数!**

(2) 另一个事实是 $g(f(x)) = x$ 对于所有的 x 都成立. 现在取一个数 x, 将 b 提升为其 x 次幂, 然后取底数为 b 的对数. 再一次地, 得到原始的数 x. 这就好像是, 取一个正数, 先平方然后再取平方根, 结果会得到原始的数. 由于 $f(x) = b^x$ 且

$g(x) = \log_b(x)$, 方程 $g(f(x)) = x$ 变为

$$\boxed{\log_b(b^x) = x} \quad (\text{对于任意的实数 } x \text{ 及 } b > 1).$$

例如对于上一节中看到的方程 $2^x = 7$, 可以对方程两边取 $\log_2$, 得到

$$\log_2(2^x) = \log_2(7).$$

等号左边正好就是 x, 因为**指数的对数就是原始的数** (前提是底数相同!). 我们再来快速地看一个例子：求

$$3^{x^2-1} = 19.$$

简单对方程两边取 $\log_3$, 得到

$$\log_3(3^{x^2-1}) = \log_3(19).$$

左边恰好就是 x^2-1, 因此我们有 $x^2-1 = \log_3(19)$. 这意味着 $x = \pm\sqrt{\log_3(19)+1}$.

9.1.4 对数法则

前述 9.1.1 节中的所有指数法则都有相应的对数版本, 它们 (毫不奇怪地) 被称为对数法则. 实际上还有额外的一条对数法则 —— 换底法则 (参见下面的法则 6), 它没有相应的指数法.[1] 因此, 下面是对于任意的底数 $b > 1$ 和正的实数 x 与 y 有效的法则：

(1) $\boxed{\log_b(1) = 0.}$

(2) $\boxed{\log_b(b) = 1.}$

(3) $\boxed{\log_b(xy) = \log_b(x) + \log_b(y).}$ **乘积的对数是对数的和.**

(4) $\boxed{\log_b(x/y) = \log_b(x) - \log_b(y).}$ **商的对数是对数的差**.

(5) $\boxed{\log_b(x^y) = y\log_b(x).}$ **对数将指数移至对数之前**. 在该方程中, y 可以是任意的实数 (正的、负的或零).

(6) **换底法则**：对于任意的底数 $b > 1$ 和 $c > 1$ 及任意的数 $x > 0$,

$$\boxed{\log_b(x) = \frac{\log_c(x)}{\log_c(b)}.}$$

这意味着, 所有的不同底数的对数函数其实是互为常数倍的. 确实, 上述方程说明

$$\log_b(x) = K\log_c(x),$$

其中 K 是常数 (它恰好等于 $1/\log_c(b)$). 当我说 “常数” 时, 我的意思是, 它不依赖于 x. 我们进而可以推出结论, $y = \log_b(x)$ 和 $y = \log_c(x)$ 的图像非常相似 —— 你

[1] 实际上, 对于指数也有一个换底法则：对于 $b > 0$, $c > 1$ 及 $x > 0$, 有 $b^x = c^{x\log_c(b)}$. 由于这里面涉及对数, 所以一般情况下, 它不被列入指数法则列表中.

只需将第二个函数的图像垂直拉伸 K 倍就能得到第一个函数的图像.



现在, 我们来看看为什么这些法则成立. 如果你想, 可以跳到下一节, 但请相信我, 如果继续阅读, 你会对对数函数有更好的理解. 不管怎样, 法则 1 相当简单: 由于对于任意的底数 $b>1$, $b^0=1$, 所以有 $\log_b(1)=0$. 同理可得法则 2: 由于对于任意的底数 $b>1$, $b^1=b$, 所以可以写出 $\log_b(b)=1$.

法则 3 有点难度. 我们必须证明 $\log_b(xy)=\log_b(x)+\log_b(y)$, 其中 x 和 y 是正的且 $b>1$. 让我们从之前已经多次注意到的一个重要事实开始 (用 A 替换之前的变量): 对于任意的 $A>0$,

$$b^{\log_b(A)}=A.$$

如果用 x、y 及 xy 分别替换 A, 则分别得到

$$b^{\log_b(x)}=x,\quad b^{\log_b(y)}=y,\quad b^{\log_b(xy)}=xy.$$

现在, 将第一个和第二个方程相乘, 然后和第三个方程相比较, 得到

$$b^{\log_b(x)}b^{\log_b(y)}=xy=b^{\log_b(xy)}.$$

那又怎么样呢? 好吧, 在左边使用指数法则 3; 由于我们必须把指数相加, 故方程变为

$$b^{\log_b(x)+\log_b(y)}=b^{\log_b(xy)}.$$

现在, 在方程两边取以 b 为底的对数来去掉底数 b; 这样, 我们就得到了对数法则 $\log_b(x)+\log_b(y)=\log_b(xy)$. 这真是不错!



至于法则 4, 我将它留给你来证明, 证明过程几乎和我们刚刚证明的法则 3 是一样的. 因此, 让我们来看看法则 5. 我们想要证明 $\log_b(x^y)=y\log_b(x)$, 其中 $x>0$, $b>1$ 且 y 是任意实数. 为了求证, 还是由前面的那个重要事实开始, 但这次用 x^y 替换 A, 得到

$$b^{\log_b(x^y)}=x^y.$$

这给了我们以一个奇怪的方式来表达 x^y. 我们也可以用 x 替换 A, 得到

$$b^{\log_b(x)}=x,$$

然后将两边提升为其 y 次幂:

$$(b^{\log_b(x)})^y=x^y.$$

等号左边正是 $b^{y\log_b(x)}$ (参见 9.1.1 节的指数法则 5). 因此, 对于 x^y, 我们有两种不同的表达, 它们必须相等:

$$b^{\log_b(x^y)}=b^{y\log_b(x)}.$$

再次对方程两边取底数为 b 的对数, 上式便简化为对数法则

$$\log_b(x^y)=y\log_b(x).$$

最后, 我们只需要证明换底法则了. 实际上要证明的是

$$\log_b(x)\log_c(b) = \log_c(x).$$

你看, 如果它是成立的, 那么在等号两边同除以 $\log_c(b)$, 便会得到法则 6. 不管怎样, 我们取上述方程, 并将左右两边分别作为 c 的次幂, 相应地得到

$$c^{\log_b(x)\log_c(b)} \quad 和 \quad c^{\log_c(x)}.$$

右边很简单：根据我们的重要事实, 它就是 x. 但左边呢？我们再次巧妙地使用指数法则 5, 得到

$$c^{\log_b(x)\log_c(b)} = c^{\log_c(b)\times\log_b(x)} = \left(c^{\log_c(b)}\right)^{\log_b(x)}.$$

由我们的重要事实 (运用两次) 可知, $c^{\log_c(b)} = b$, $b^{\log_b(x)} = x$, 于是推导出结论

$$c^{\log_b(x)\log_c(b)} = \left(c^{\log_c(b)}\right)^{\log_b(x)} = b^{\log_b(x)} = x.$$

因此, 上述这两个量

$$c^{\log_b(x)\log_c(b)} \quad 和 \quad c^{\log_c(x)}$$

都可简化为 x! 它们必须相等, 因此如果去掉底数 c (两边取以 c 为底的对数), 我们就能得到想要的方程

$$\log_b(x)\log_c(b) = \log_c(x).$$

如果你选择了努力去理解所有这些证明, 那你做得真是不错.

9.2 e 的定义

到目前为止, 我们还没有做过涉及指数函数或对数函数的任何微积分运算. 现在就开始做些吧. 我们会先从极限开始, 然后进入导数. 在这一过程中, 需要引入一个新的常数 e. 和 π 一样, 它也是一个特别的数 —— 当你对数学探索得足够深时, 它就会不知从哪儿冒出来. 一种探索 e 从何而来的方法会涉及一点金融问题.

9.2.1 一个有关复利的问题

很久以前, 一个名叫伯努利的家伙回答了一个有关复利的问题. 下面就是该问题. 设想你在一家银行有一个银行账户, 该银行付给你一个慷慨的利息年利率 12%, 一年计一次复利. 你将一笔初始存款存入账户. 每一年你的财富增加 12%. 这意味着, n 年后, 你的财富会增加到原来的 $(1+0.12)^n$ 倍. 特别地, 一年后, 你的财富就是 $(1+0.12)$ 乘以原始存款. 如果你最开始存入了 100 美元, 年底你会得到 112 美元.

现在设想你发现另一家银行, 它也提供 12% 的年利率, 但它是一年计两次复利. 当然, 每半年, 你不会得到 12%; 你必须用它除以 2. 简单说, 这意味着, 每六个月你会得到 6% 的利息. 因此, 如果你将钱存入这个银行账户, 那么一年后, 它会以 6% 的利息计算复利两次; 结果就是你的财富会增加到原来的 $(1+0.06)^2$ 倍, 其结果 1.1236. 因此, 如果你最开始存入了 100 美元, 年底你会得到 112.36 美元.

第二个账户的收益比第一个略好一些. 稍作思考, 你不难发现这很说得通——复利是有益的, 因此在相同的年利率下, 复利计算得越频繁, 结果就会越好. 我们来试着计算一下年利率为 12%、每年计三次复利. 我们取 12%, 并将它除以 3 会得到 4%, 然后计算复利三次, 我们的财富将会增加到原来的 $(1+0.04)^3$ 倍, 其结果是 1.124 864. 这还是高了些. 要是每年计四次呢？那将是 $(1+0.03)^4$ 倍, 结果近似为 1.1255, 这就更高了. 现在的问题是, 它会止于何处？如果你以相同的年利率计算复利越来越频繁, 一年后你会得到大把大把的现金, 还是这一切有某种上限？

9.2.2　问题的答案

为了回答这个问题, 让我们来求助于一些符号. 首先, 假设以年利率 12% 每年计 n 次复利. 这意味着, 每次计算复利时, 复利的利率是 $0.12/n$. 在一年中计算 n 次后, 我们的原始财富增长的倍数为

$$\left(1+\frac{0.12}{n}\right)^n.$$

我们想要知道, 如果复利计算得越来越频繁时会怎样; 实际上, 这意味着我们允许 n 变得越来越大. 也就是说, 我们想知道, 当 $n\to\infty$ 时上式的极限

$$\lim_{n\to\infty}\left(1+\frac{0.12}{n}\right)^n$$

到底是什么？我们也很想知道, 当利率不是12% 时会发生什么. 因此, 用 r 代替 0.12, 并关心更一般的极限

$$L=\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n.$$

如果结果表明该极限 (我称之为 L) 是无限的, 那么通过越来越频繁地计算复利, 你在一年中可以得到越来越多的钱. 另一方面, 如果结果表明它是有限的, 我们就必须得出结论, 在一个年利率 r 的情况下, 不管复利计算得多么频繁, 我们的财富的增长幅度都是有上限的. 这类似于一种 “速率极限”, 或更精确地说, 一种 “财富增长极限”. 给定一个固定的年利率 r 以及一年的时间, 不管复利计算得多么频繁, 你都不可能让财富的增长超过上述极限的值 (假设它是有限的).

出现在极限中的量 $(1+r/n)^n$ 是复利计算公式的一个特例. 一般地, 假设你以现金 A 美元开始, 并且将它存入一个银行账户, 年利率为 r, 一年计 n 次复利, 那么在 t 年中, 将以每次 r/n 的利率计算 nt 次复利. 因此, t 年后, 你的财富由以下公式给出：

$$\text{年利率为 } r\text{, 一年计 } n \text{ 次复利, } t \text{ 年后的财富} = A\left(1+\frac{r}{n}\right)^{nt}.$$

我们不妨从 1 美元 (故 $A=1$) 开始, 来看看一年 (故 $t=1$) 后会发生什么, 然后看

看如果我们在一年中复利计算得越来越频繁时极限会怎样.

现在来计算极限

$$L = \lim_{n \to \infty} \left(1 + \frac{r}{n}\right)^n.$$

首先, 设 $h = r/n$, 这样 $n = r/h$. 那么当 $n \to \infty$ 时, 我们看到 $h \to 0^+$ (因为 r 是常数), 故

$$L = \lim_{h \to 0^+} (1+h)^{r/h}.$$

现在可以使用指数法则将之写成

$$L = \lim_{h \to 0^+} ((1+h)^{1/h})^r.$$

然后来变个魔术, 设

$$\mathrm{e} = \lim_{h \to 0^+} (1+h)^{1/h}.$$

这里的危险在哪里呢? 好吧, 这个极限有可能不存在. 还好事实表明, 它存在. 如果你想要知道原因的话, 请参见附录 A 的 A.5 节. 不管怎样, 我们有了一个特殊的数 e, 关于其更多细节我们马上就会看到. 不过还是先回到我们的极限. 现在有

$$L = \lim_{h \to 0^+} ((1+h)^{1/h})^r = \mathrm{e}^r.$$

这就是我们要找的答案! 将以上所有的步骤综合在一起, 可以看到整个运算是怎样推进的. 由于 $h = r/n$, 有

$$L = \lim_{n \to \infty} \left(1 + \frac{r}{n}\right)^n = \lim_{h \to 0^+} (1+h)^{r/h} = \lim_{h \to 0^+} ((1+h)^{1/h})^r = \mathrm{e}^r.$$

这意味着, 在年利率 r、复利计算得越来越频繁的情况下, 你的财富会增长到一个非常接近于 e^r 的量, 但绝不会超过它. 量 e^r 就是我们要找的 "财富增长极限". 得到这个增长倍数的唯一途径就是连续地计复利 —— 也就是说, 每时每刻都在计复利!

因此, 假设你由 A 美元的现金开始, 并将它存入一个银行账户, 它以年利率 r 连续计复利. 这样一年后, 你会有 $A\mathrm{e}^r$ 美元. 两年后, 你会有 $A\mathrm{e}^r \times \mathrm{e}^r = A\mathrm{e}^{2r}$ 美元. 我们很容易一直重复这个过程, 并看到 t 年后, 你会有 $A\mathrm{e}^{rt}$ 美元. 由于指数法则, 这实际上对于分数年也成立. 因此, 由 A 美元开始,

$$\boxed{\text{以年利率 } r \text{ 连续计复利, } t \text{ 年后的财富} = A\mathrm{e}^{rt}.}$$

比较该公式和 $r = A\left(1 + \dfrac{r}{n}\right)^{nt}$. 量 $A(1+r/n)^{nt}$ 和 $A\mathrm{e}^{rt}$ 看起来很不同, 但对于很大的 n, 它们几乎是一样的.

9.2.3　更多关于 e 和对数函数的内容

让我们来更深入地看一下数 e 吧. 记得

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = \mathrm{e}^r,$$

我们可以用 1 替换 r, 得到

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \mathrm{e}.$$

当然, $r=1$ 对应于一个 100% 的年利率. 让我们列一个 $(1+1/n)^n$ 的值的表, 对于不同的 n 值, 结果保留三位小数：

n	1	2	3	4	5	10	100	1000	10 000	100 000
$\left(1+\frac{1}{n}\right)^n$	2	2.25	2.353	2.441	2.488	2.594	2.705	2.717	2.718	2.718

即使只是一年计一次复利, 这个极高的利率也可以使你的钱一年后翻倍 (也就是第二列下面一行中的 “2”). 尽管如此, 看上去我们仍然不可能做得比 2.718 更好, 即使每一年计复利很多很多次. 我们的数 e, 也就是上表中第二行的数在 $n\to\infty$ 时的极限, 事实证明是一个无理数, 其小数展开式前几位如下：

$$\mathrm{e} = 2.718\,281\,828\,459\,045\,23\cdots$$

似乎在开头部分存在一个模式, “1828” 有重复出现, 但这只是个巧合. 在实践中, 知道 e 比 2.7 大一点就已经足够了.

现在, 如果 $x=\mathrm{e}^r$, 那么 $r=\log_{\mathrm{e}}(x)$. 事实表明, 取以 e 为底的对数是如此常见, 以至于我们甚至可以将它用另一种方式写出：$\ln(x)$, 而不是 $\log_{\mathrm{e}}(x)$. 表达式 “$\ln(x)$” 不读作 “lin x” 或其他诸如此类, 而是可以读作 “log x”, 或 “ell en x”, 又或特别严谨地 ——“x 的自然对数”. 事实上, 大多数数学家写不带底数的 $\log(x)$ 来表示和 $\log_{\mathrm{e}}(x)$ 及 $\ln(x)$ 相同的意思. 底数为 e 的对数称为**自然对数**. 在下一节, 当对 $\log_b(x)$ 关于 x 求导时, 我们会看到为什么说它 “自然” 的一个原因.

我们有了一个新的底数 e, 以及以 e 为底时的一个新的对数写法, 再来看看迄今已经看到的对数法则和公式吧. 看你是否能让自己确信, 对于 $x>0$ 和 $y>0$, 下列公式都成立：

$$\boxed{\mathrm{e}^{\ln(x)}=x}\quad \boxed{\ln(\mathrm{e}^x)=x}\quad \boxed{\ln(1)=0}\quad \boxed{\ln(\mathrm{e})=1}$$

$$\boxed{\ln(xy)=\ln(x)+\ln(y)}\quad \boxed{\ln\left(\frac{x}{y}\right)=\ln(x)-\ln(y)}\quad \boxed{\ln(x^y)=y\ln(x)}$$

(事实上, 在第二个公式中, x 甚至可以是负数或 0; 在最后一个公式中, y 可以是负数或 0.) 不管怎样, 知道在这种形式下的对数法则是很值得的, 因为从现在起, 我

们会几乎总与自然对数打交道.



在我们讨论对数函数和指数函数求导之前, 再多看一点. 假设你取重要极限

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = \mathrm{e}^r,$$

这一次, 替换 $h=1/n$. 正如我们在上一节注意到的, 当 $n\to\infty$ 时, 我们有 $h\to 0^+$. 因此, 用 $1/h$ 替换 n, 得到

$$\lim_{h\to 0^+}(1+rh)^{1/h} = \mathrm{e}^r.$$

这是一个右极限. 事实上, 你可以用 $h\to 0$ 替换 $h\to 0^+$, 对于双侧极限仍然成立. 我们所需的就是证明, 左极限是 e^r, 然后左极限等于右极限, 故双侧极限也等于 e^r. 为此, 考虑

$$\lim_{h\to 0^-}(1+rh)^{1/h} = ?$$

用 $-t$ 替换 h, 那么当 $h\to 0^-$ 时, $t\to 0^+$. (当 h 是一个很小的负数时, $t=-h$ 就是一个很小的正数.) 故

$$\lim_{h\to 0^-}(1+rh)^{1/h} = \lim_{t\to 0^+}(1-rt)^{-1/t}.$$

由于对于任意的 $A\neq 0$, $A^{-1}=1/A$, 我们可以重新将极限写成

$$\lim_{t\to 0^+}\frac{1}{(1+(-r)t)^{1/t}}.$$

分母就是利率为 $-r$ 而不是 r 的经典极限. 这意味着, 当 $t\to 0^+$ 时, 在该极限中, 分母趋于 e^{-r}. 因此, 综合起来有

$$\lim_{h\to 0^-}(1+rh)^{1/h} = \lim_{t\to 0^+}(1-rt)^{-1/t} = \lim_{t\to 0^+}\frac{1}{(1+(-r)t)^{1/t}} = \frac{1}{\mathrm{e}^{-r}} = \mathrm{e}^r.$$

由于 $\mathrm{e}^{-r}=1/\mathrm{e}^r$, 故最后一步成立. 这样, 我们完成了想要的证明. 让我们在所有的公式中将 r 改为 x(为什么不呢?) 并总结已经发现的如下事实:

$$\boxed{\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \mathrm{e}^x} \quad 和 \quad \boxed{\lim_{h\to 0}(1+xh)^{1/h} = \mathrm{e}^x.}$$

当 $x=1$ 时, 我们得到 e 的两个公式:

$$\boxed{\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \mathrm{e}} \quad 和 \quad \boxed{\lim_{h\to 0}(1+h)^{1/h} = \mathrm{e}.}$$

这些公式非常重要! 在下面的 9.4.1 节中, 我们将看到一些如何使用它们的例子. 马上, 我们也会使用其中之一来对对数函数求导.

9.3　对数函数和指数函数求导

现在情形变复杂了. 令 $g(x) = \log_b(x)$. g 的导数是什么呢? 使用导数的定义, 我们得到

$$g'(x) = \lim_{h \to 0} \frac{g(x+h) - g(x)}{h} = \lim_{h \to 0} \frac{\log_b(x+h) - \log_b(x)}{h}.$$

但如何来化简这个杂乱的公式呢? 当然是使用对数法则! 首先, 使用 9.1.4 节中的法则 4, 将对数的差转化为对数的商:

$$g'(x) = \lim_{h \to 0} \frac{1}{h} \log_b\left(\frac{x+h}{x}\right).$$

我们可以将分式化简为 $(1 + h/x)$, 并使用对数法则 5, 将因子 $1/h$ 提至指数的位置. 故

$$g'(x) = \lim_{h \to 0} \log_b\left(1 + \frac{h}{x}\right)^{1/h}.$$

现在让我们暂时忘记 $\log_b$. 当 h 趋于 0 时,

$$\left(1 + \frac{h}{x}\right)^{1/h}$$

会怎样呢? 也就是说,

$$\lim_{h \to 0} \left(1 + \frac{h}{x}\right)^{1/h}$$

是什么呢? 在上一节中, 我们看到了

$$\lim_{h \to 0} (1 + hr)^{1/h} = \mathrm{e}^r;$$

因此, 如果用 $1/x$ 替换 r, 就会有

$$\lim_{h \to 0} \left(1 + \frac{h}{x}\right)^{1/h} = \mathrm{e}^{1/x}.$$

所以如果回到 $g'(x)$ 的表达式, 我们会看到

$$g'(x) = \lim_{h \to 0} \log_b\left(1 + \frac{h}{x}\right)^{1/h} = \log_b(\mathrm{e}^{1/x}).$$

事实上, 我们甚至可以再次使用对数法则 5 将表达式进一步化简 —— 将指数 $1/x$ 提至对数符号之前, 这样就证明了

$$\frac{\mathrm{d}}{\mathrm{d}x} \log_b(x) = \frac{1}{x} \log_b(\mathrm{e}).$$

现在, 设 $b = \mathrm{e}$, 这样就能求以 e 为底的对数的导数了, 得到

$$\frac{\mathrm{d}}{\mathrm{d}x} \log_\mathrm{e}(x) = \frac{1}{x} \log_\mathrm{e}(\mathrm{e}).$$

但等一下 —— 根据对数法则 2, $\log_e(e)$ 等于 1. 因此, 这意味着

$$\frac{d}{dx}\log_e(x)=\frac{1}{x}.$$

这相当好. 实际上这非常非常好. 着实可以说是迷人. 谁会想到 $\log_e(x)$ 的导数就是 $1/x$ 呢? 这正是为什么以 e 为底的对数被称为自然对数的原因之一. 我们将 $\log_e(x)$ 写作 $\ln(x)$(在上一节我们给出了这个定义), 得到重要公式

$$\boxed{\frac{d}{dx}\ln(x)=\frac{1}{x}.}$$

此外, 上面的 $\log_b(x)$ 的导数的表达式 $\frac{1}{x}\log_b(e)$ 可以通过换底法则 (就是 9.1.4 节中的法则 6) 用自然对数写出. 你看, 通过将底换为 e, 得到

$$\log_b(e)=\frac{\log_e(e)}{\log_e(b)}=\frac{1}{\ln(b)}.$$

因此, 有

$$\boxed{\frac{d}{dx}\log_b(x)=\frac{1}{x\ln(b)}.}$$

这是表达一个不是以 e 为底的对数的导数的最好方式了. 现在再来看看这个: 如果 $y=b^x$, 那么我们知道 $x=\log_b(y)$. 现在对其关于 y 求导. 使用上述公式并用 y 替换 x, 得到

$$\frac{dx}{dy}=\frac{1}{y\ln(b)}.$$

根据链式求导法则, 可以上下颠倒得到

$$\frac{dy}{dx}=y\ln(b).$$

由于 $y=b^x$, 我们就证明了下面这个很好的公式:

$$\boxed{\frac{d}{dx}(b^x)=b^x\ln(b).}$$

特别是, 如果 $b=e$, 那么 $\ln(b)=\ln(e)=1$. (这是另一种形式的对数法则 2. 回想一下, $\ln(e)=\log_e(e)=1$.) 因此, 如果 $b=e$, 公式变为

$$\boxed{\frac{d}{dx}(e^x)=e^x.}$$

这是一个相当奇怪的公式. 如果 $h(x)=e^x$, 那么也有 $h'(x)=e^x$ (函数 h 是它自身的导数!). 当然, e^x 的 (关于 x 的) 二阶导还是 e^x, 三阶导、四阶导等也是如此.

指数函数和对数函数求导的例子

现在来看一下如何应用上述公式吧. 首先, 如果 $y=e^{-3x}$, 那么 dy/dx 是什么?


如果设 $u=-3x$, 那么 $y=\mathrm{e}^u$. 我们有

$$\frac{\mathrm{d}y}{\mathrm{d}u}=\frac{\mathrm{d}}{\mathrm{d}u}(\mathrm{e}^u)=\mathrm{e}^u \quad 和 \quad \frac{\mathrm{d}u}{\mathrm{d}x}=\frac{\mathrm{d}}{\mathrm{d}x}(-3x)=-3.$$

根据链式求导法则,

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}=\mathrm{e}^u(-3)=-3\mathrm{e}^{-3x};$$

注意到最后一步用 $-3x$ 替换了 u. 事实上, 这是另一个很好的法则的一个特例：如果 a 是常数, 那么

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{ax}=a\mathrm{e}^{ax}.$$


通过设 $u=ax$, 我们可以用同样的方法证明此公式. 实际上, 它和我们在 7.2.1 节结尾部分看到的原理是一样的：**如果用 ax 替换 x, 那么当你求导的时候, 在最前面会有一个额外的因子 a**. 因此, 对于例如 $\ln(8x)$ 关于 x 求导就应该不成问题. 事实上,

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln(8x))=8\times\frac{1}{8x},$$

因为 $\ln(8x)$ 关于 x 的导数是 $1/x$. 现在, 消去因子 8, 看到

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln(8x))=8\times\frac{1}{8x}=\frac{1}{x}.$$

这真是奇怪 ——$\ln(8x)$ 的导数和 $\ln(x)$ 的导数是一样的! 但稍作思考, 你就不会觉得它有那么奇怪了：由于 $\ln(8x)=\ln(8)+\ln(x)$, 量 $\ln(8x)$ 和 $\ln(x)$ 实际上只相差一个常数, 故关于 x 它们有相同的导数.


下面是一个难一点的例子：

$$如果\ y=\mathrm{e}^{x^2}\log_3(5^x-\sin(x)),\quad \frac{\mathrm{d}y}{\mathrm{d}x}\ 是什么?$$

让我们来使用乘积法则和链式求导法则. 设 $u=\mathrm{e}^{x^2}$, $v=\log_3(5^x-\sin(x))$, 故 $y=uv$. 对于乘积法则, 需要对 u 和 v (关于 x) 求导, 因此让我们一个一个地进行. 我们从 $u=\mathrm{e}^{x^2}$ 开始, 设 $t=x^2$, 因此 $u=\mathrm{e}^t$; 然后, 使用链式求导法则, 有

$$\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{\mathrm{d}u}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x}=\mathrm{e}^t(2x)=2x\mathrm{e}^{x^2}.$$

至于 v, 令 $s=5^x-\sin(x)$, 于是 $v=\log_3(s)$. 根据链式求导法则,

$$\frac{\mathrm{d}v}{\mathrm{d}x}=\frac{\mathrm{d}v}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}x}=\frac{1}{s\ln(3)}(5^x\ln(5)-\cos(x))=\frac{5^x\ln(5)-\cos(x)}{\ln(3)(5^x-\sin(x))}.$$

这里使用了上一节中 $\log_b(x)$ (现在 $b=3$) 和 b^x(现在 $b=5$) 的导数公式. 不管怎样, 由于 $y=uv$, 有

$$\frac{\mathrm{d}y}{\mathrm{d}x}=v\frac{\mathrm{d}u}{\mathrm{d}x}+u\frac{\mathrm{d}v}{\mathrm{d}x}=\log_3(5^x-\sin(x))2x\mathrm{e}^{x^2}+\mathrm{e}^{x^2}\frac{5^x\ln(5)-\cos(x)}{\ln(3)(5^x-\sin(x))}.$$

像往常一样, 这有些杂乱, 但这个例子确实很好地说明了一点：只要你知道指数函数和对数函数求导的基本公式 (也就是上一节中的方框公式), 那么相关求解就完全不成问题了.

9.4 求解指数函数或对数函数的极限

现在是时候来看看如何求解一些极限问题了. 正如在我们之前所见的所有求极限问题中那样, 非常重要的一点是, 注意到你是在哪里计算函数的极限的：是在 0 附近 (也就是说, 小的数), 还是在 ∞ 或 $-\infty$ 附近 (也就是说, 大的数), 又或者在某个既不大也不小的数附近？我们将就这些情况中的一些分别对指数函数和对数函数进行略微深入的讨论. 让我们先从涉及 e 的定义的极限开始.

9.4.1 涉及 e 的定义的极限

考虑极限

$$\lim_{h\to 0}(1+3h^2)^{1/3h^2}.$$

它看上去和 9.2.3 节中涉及 e 的极限

$$\lim_{h\to 0}(1+h)^{1/h}=\mathrm{e}$$

非常相似. 如果我们采用这个极限, 并一律用 $3h^2$ 代替 h, 那么会得到

$$\lim_{3h^2\to 0}(1+3h^2)^{1/3h^2}=\mathrm{e}.$$

这几乎就是我们想要的. 我们所需做的只是注意到, 当 $h\to 0$ 时, $3h^2\to 0$, 故

$$\lim_{h\to 0}(1+3h^2)^{1/3h^2}=\mathrm{e}.$$

同理, 我们可以证明 (例如)

$$\lim_{h\to 0}(1+\sin(h))^{1/\sin(h)}=\mathrm{e}.$$

确实, 如果用任意的当 $h\to 0$ 时自身趋于 0 的量替换 h, 就像 $3h^2$ 或 $\sin(h)$, 则极限仍是 e. 那么

$$\lim_{h\to 0}(1+\cos(h))^{1/\cos(h)}$$

又怎样呢？由于当 $h\to 0$ 时, $\cos(h)\to 1$, 因此你不能照搬之前的论证. 事实上, 如果将 $h=0$ 代入到表达式 $(1+\cos(h))^{1/\cos(h)}$ 中, 那么会得到 $(1+1)^1=2$, 故上述极限实际上等于 2.

现在考虑

$$\lim_{h\to 0}(1+h^2)^{1/3h^2}.$$

h^2 和 $3h^2$ 这两项不匹配. 它们很相似, 但系数不同. 为此, 需要将指数 $1/3h^2$ 写作 $(1/h^2) \times (1/3)$, 并使用指数法则：

$$\lim_{h\to0}(1+h^2)^{1/3h^2} = \lim_{h\to0}(1+h^2)^{(1/h^2)\times(1/3)} = \lim_{h\to0}\left((1+h^2)^{1/h^2}\right)^{1/3}.$$

由于两个 h^2 相匹配, 故大括号中的部分趋于 e, 而整个的极限是 $\mathrm{e}^{1/3}$.

下面是一个略难一些的例子：极限

$$\lim_{h\to0}(1-5h^3)^{2/h^3}$$

是什么？这里恼人的是, 量 $-5h^3$ 和 h^3 不十分匹配, 并且那里还有一个 2. 我们需要改动指数 $2/h^3$ 以便它和 $-5h^3$ 相匹配. 最好的方法是, 先注意到完美的形式应该是

$$\lim_{h\to0}(1-5h^3)^{1/(-5h^3)},$$

因为该极限就是 e. 这时两个 $-5h^3$ 相匹配, 因此它只是用 $-5h^3$ 代替了 h 的经典极限

$$\lim_{h\to0}(1+h)^{1/h} = \mathrm{e},$$

但不幸的是, 我们还需要再多做一些工作. 我们需要将 $1/(-5h^3)$ 变为 $2/h^3$. 为了实现这一变化, 必须用 -5 与之相乘来消去分母中的 -5, 然后再用 2 与之相乘来修正分子. 总的效果就是应该用 -10 与之相乘. 这样得到

$$\begin{aligned}\lim_{h\to0}(1-5h^3)^{2/h^3} &= \lim_{h\to0}(1-5h^3)^{(1/(-5h^3))\times(-10)}\\ &= \lim_{h\to0}\left((1-5h^3)^{1/(-5h^3)}\right)^{-10} = \mathrm{e}^{-10}.\end{aligned}$$

9.4.2　指数函数在 0 附近的行为

我们想要理解, 当 x 非常接近于 0 时, e^x 的行为会如何. 事实上, 由于 $\mathrm{e}^0 = 1$, 我们知道

$$\lim_{x\to0}\mathrm{e}^x = \mathrm{e}^0 = 1.$$

当然, 可以用任意的当 $x \to 0$ 时自身趋于 0 的量来替换 x, 来得到相同的极限. 例如,

$$\lim_{x\to0}\mathrm{e}^{x^2} = \mathrm{e}^{0^2} = 1.$$

因此, 求

$$\lim_{x\to0}\frac{\mathrm{e}^{x^2}\sin(x)}{x}$$

的方法是, 将上式进行如下拆分：

$$\lim_{x\to 0}\frac{e^{x^2}\sin(x)}{x}=\lim_{x\to 0}(e^{x^2})\left(\frac{\sin(x)}{x}\right).$$

当 $x\to 0$ 时, 两个因子都趋于 1, 故整个极限为 $1\times 1=1$. 下面则是一个更难求解的例子：



$$\lim_{x\to\infty}\frac{2x^2+3x-1}{e^{1/x}(x^2-7)}.$$

当 x 变得非常大时, $1/x$ 会变得非常接近于 0, 故 $e^{1/x}$ 非常接近于 1 并可被忽略. 所以你最好将以上极限重写为

$$\lim_{x\to\infty}\frac{1}{e^{1/x}}\times\frac{2x^2+3x-1}{x^2-7}.$$

第一个分式趋于 1, 而使用 4.3 节的技巧, 可以证明第二个因子趋于 2, 故极限是 2.

这种方法在指数项出现在一个乘积或商当中时最好用, 但对于诸如

$$\lim_{h\to 0}\frac{e^h-1}{h}$$

这样的形式就彻底无能为力了. 你可能想用 1 替换 e^h, 看上去没错, 但你会得到一个无用的 0/0 的情况. 这里的问题是, 有一个 e^h 和 1 的差, 当 h 在 0 附近时, 它会变得非常小. 那我们应该怎么办呢? 正如我们在 6.5 节中看到的, 当虚拟变量本身在分母上时, 极限可能是一个伪装的导数. 试着设 $f(x)=e^x$, 这样 $f'(x)=e^x$ (正如我们在 9.3 节中看到的). 在这种情况下, 标准公式

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x)$$

变为

$$\lim_{h\to 0}\frac{e^{x+h}-e^x}{h}=e^x.$$

现在, 所需做的只是用 0 替换 x. 由于 $e^0=1$, 我们得到以下有用的事实：

$$\boxed{\lim_{h\to 0}\frac{e^h-1}{h}=1.}$$

再一次地, 可以用任意的很小的量来替换 h. 例如,

$$\lim_{s\to 0}\frac{e^{3s^5}-1}{s^5}=\lim_{s\to 0}\frac{e^{3s^5}-1}{3s^5}\times 3=1\times 3=3.$$

标准的匹配技巧再次奏效. 这实际上和我们在多项式型的极限问题 (第 4 章)、小数情况的三角函数极限问题 (第 7 章) 以及 9.4.1 节的极限问题中所用的技巧是一样的.

9.4.3 对数函数在 1 附近的行为



现在让我们来看看对数函数在 1 附近的行为会如何. 事实表明, 其行为和指数函数在 0 附近的行为十分相似. 我们知道 $\ln(1) = 0$, 但

$$\lim_{h\to 0}\frac{\ln(1+h)}{h}$$

是什么呢? 不管你是否相信, 这其实是导数伪装的极限 (参见 6.5 节) 的另一个例子. 设 $f(x) = \ln(x)$, 这样, 正如我们在 9.3 节看到的, $f'(x) = 1/x$. 现在等式

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = f'(x)$$

变为, 对于任意的 x,

$$\lim_{h\to 0}\frac{\ln(x+h)-\ln(x)}{h} = \frac{1}{x}.$$

剩下要做的只是将 $x = 1$ 代入并得到

$$\lim_{h\to 0}\frac{\ln(1+h)-\ln(1)}{h} = \frac{1}{1}.$$

由于 $\ln(1) = 0$, 上式简化为

$$\boxed{\lim_{h\to 0}\frac{\ln(1+h)}{h} = 1.}$$

再一次地, 可以用任意的当 $h \to 0$ 时自身趋于 0 的量来替换 h, 而极限仍将是 1. 例如, 为了求

$$\lim_{h\to 0}\frac{\ln(1-7h^2)}{5h^2},$$

你必须改动分母, 使它看起来像 $-7h^2$:

$$\lim_{h\to 0}\frac{\ln(1-7h^2)}{5h^2} = \lim_{h\to 0}\frac{\ln(1-7h^2)}{-7h^2}\times\frac{-7h^2}{5h^2}.$$

这不过是我们那个常用的技巧, 分子分母同时乘以一个有用的量 (在该例中是 $-7h^2$). 不管怎样, 由于两个 $-7h^2$ 相匹配, 故第一个分式的极限是 1, 而第二个分式正好化简为 $-7/5$. 因此极限就是 $-7/5$.

9.4.4 指数函数在 ∞ 或 $-\infty$ 附近的行为

现在我们想要理解, 当 $x \to \infty$ 或 $x \to -\infty$ 时, e^x 的行为会如何. 让我们再来看看 e^x 的图像吧, 如图 9-2 所示.

注意: 以上曲线看起来好像要在图像的左侧触碰到 x 轴, 但它**没有**; 回想一下, 对于所有的 x, $\mathrm{e}^x > 0$, 因此没有 x 轴截距. (这是一个说明不能太过依赖图形计算器的很好例子!) 不管怎么样, 看起来我们应该至少有

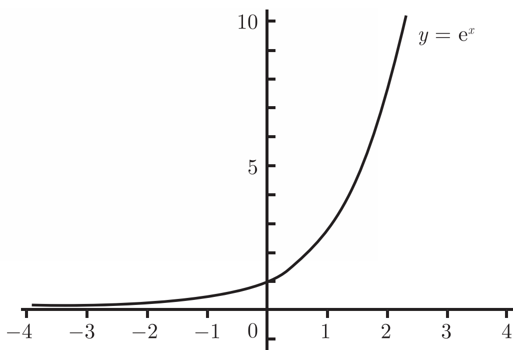


图 9-2

$$\lim_{x\to\infty} \mathrm{e}^x = \infty \quad 和 \quad \lim_{x\to-\infty} \mathrm{e}^x = 0.$$

如果用某个其他的底数替换 e 会怎样呢? 例如, 考虑

$$\lim_{x\to\infty} 2^x \quad 和 \quad \lim_{x\to\infty} \left(\frac{1}{3}\right)^x.$$

为了处理第一个极限, 需要使用等式 $A = \mathrm{e}^{\ln(A)}$, 其中 $A = 2^x$, 写出

$$2^x = \mathrm{e}^{\ln(2^x)} = \mathrm{e}^{x\ln(2)}.$$

现在, 当 $x \to \infty$ 时, 我们也有 $x\ln(2) \to \infty$, 故第一个极限是 ∞. 至于第二个极限, 可以使用相同的技巧, 写出

$$\left(\frac{1}{3}\right)^x = \frac{1}{3^x} = \frac{1}{\mathrm{e}^{x\ln(3)}}.$$

当 $x \to \infty$ 时, 我们看到 $\mathrm{e}^{x\ln(3)} \to \infty$, 故其倒数趋于 0. 这样, 就证明了

$$\lim_{x\to\infty} 2^x = \infty \quad 和 \quad \lim_{x\to\infty} \left(\frac{1}{3}\right)^x = 0.$$

可见下面这个重要极限要分几种情况:

$$\lim_{x\to\infty} r^x = \begin{cases} \infty & 如果\ r > 1, \\ 1 & 如果\ r = 1, \\ 0 & 如果\ 0 \leqslant r < 1. \end{cases}$$

当 $r = 1$ 时, 中间的情况显然成立, 因为对于所有的 $x \geqslant 0$, $1^x = 1$. 我们可以用处理前面 2^x 和 $(1/3)^x$ 的极限时相同的方法来证明其他两种情况, 即将 r^x 写作 $\mathrm{e}^{x\ln(r)}$.

这还不是故事的全部. 极限

$$\lim_{x\to\infty} \mathrm{e}^x = \infty$$

说明当 x 变大时, e^x 变得越来越大 (你想要多大就多大). 但这发生得有多快呢? 毕竟还有

$$\lim_{x\to\infty} x^2 = \infty.$$

x^2 或 e^x, 哪一个增长得更快呢? 答案是, 当 x 很大时, e^x 比 x^2 增长得更快. 毕竟, 当 $x = 100$ 时, 量 x^2 只是 100×100, 而

$$\mathrm{e}^{100} = \mathrm{e} \times \mathrm{e} \times \cdots \times \mathrm{e}.$$

后者有一百个因子 e, 而前者只有两个因子 100, 故 e^{100} 远远大于 100^2. 当 x 变得更大时, 情况变得对 e^x 更为有利. 由于 e^x 远远大于 x^2, 当你用 x^2 除以 e^x 时, 应该得到一个很小的数. 事实上,

$$\lim_{x\to\infty} \frac{x^2}{\mathrm{e}^x} = 0.$$

我们要到第 14 章看过洛必达法则后再来证明上式. 现在我只想指出, 如果你用 x 的任意次幂替换 x^2, 上述极限依然成立. 就连 x^{999} 都不能和 e^x 抗衡. 当 x 是十亿时, x^{999} 是十亿 999 次重复相乘的结果, 但 e^x 是 e 十亿次重复相乘的结果! 尽管 e 比十亿小很多, 但当 x 很大时, e^x 会让 x^{999} 相形见绌. 因此, 一般来说, 我们有以下原则:

指数函数增长迅速: 不管 n 有多大,

$$\boxed{\lim_{x\to\infty} \frac{x^n}{\mathrm{e}^x} = 0.}$$

事实上, 对上式进行一些微调, 可以得到一个更一般的陈述:

$$\lim_{x\to\infty} \frac{\text{多项式型}}{\text{指数为大的、正的多项式型的指数函数}} = 0.$$

例如,

$$\lim_{x\to\infty} \frac{x^8 + 100x^7 - 4}{\mathrm{e}^x} = 0.$$

为了证明这一点, 我们可以简单地将分式分成三部分, 每一部分都趋于 0, 因为指数函数增长迅速. 不那么明显的是,

$$\lim_{x\to\infty} \frac{x^{10\,000} + 300x^9 + 32}{\mathrm{e}^{2x^3-19x^2-100}} = 0.$$

这里的关键是, 当 x 很大时, $2x^3 - 19x^2 - 100$ 表现得就像 $2x^3$, 因此指数确实是大的、正的多项式型.[①] 事实上, 我们可以用任意的大于 1 的底数来替换 e. 例如,

$$\lim_{x\to\infty} \frac{x^{10\,000} + 300x^9 + 32}{2^{2x^3-19x^2-100}} = 0$$

① 如果你真想让这板上钉钉的话, 就必须写出诸如这样巧妙的关系, 对于足够大的 x, $2x^3-19x^2-100 > x^3$. 毕竟, 如果 $2x^3 - 19x^2 - 100$ 表现得像 $2x^3$, 那么很明显, 它也会大于 x^3. 因此, 分母大于 e^{x^3}. 现在用 u 替换 x^3, 这样分母就是 e^u, 而分子是某个很容易处理的表达式. 最后, 使用三明治定理.

另一个变形涉及这样一个事实, e^{-x} 是 $1/\mathrm{e}^x$ 的另一种写法. 下面是一个有关的例子:

$$\lim_{x\to\infty}(x^5+3)^{101}\mathrm{e}^{-x}.$$

我们可以将它写成

$$\lim_{x\to\infty}(x^5+3)^{101}\mathrm{e}^{-x}=\lim_{x\to\infty}\frac{(x^5+3)^{101}}{\mathrm{e}^x}=0;$$

这里的极限是 0, 因为指数函数增长迅速. 现在, 考虑一个与之非常相似的极限

$$\lim_{x\to-\infty}(x^5+3)^{101}\mathrm{e}^{x}.$$

这当然涉及了 e^x 在 $-\infty$ 附近的行为, 但通过设 $t=-x$, 我们可以将情形转换 $+\infty$. 可以看到, 当 $x\to-\infty$ 时, 有 $t\to+\infty$. 因此,

$$\begin{aligned}\lim_{x\to-\infty}(x^5+3)^{101}\mathrm{e}^{x}&=\lim_{t\to\infty}((-t)^5+3)^{101}\mathrm{e}^{-t}\\&=\lim_{t\to\infty}\frac{(-t^5+3)^{101}}{\mathrm{e}^t}=0.\end{aligned}$$

再一次地, 极限是 0, 因为分子是一个多项式 (其首项为负, 但这并不要紧). 因此, 通过做替换 $t=-x$, 你可以处理当 $x\to-\infty$ 时的 e^x 的极限; 这也意味着, 现在你必须处理当 $t\to\infty$ 时 e^{-t} 的极限, 不过这只需将 e^{-t} 写成 $1/\mathrm{e}^t$.

9.4.5 对数函数在 ∞ 附近的行为

旅程继续. 现在让我们来看看当 x 是一个大的正数时 $\ln(x)$ 的行为会如何. (回想一下, 你不能取任何负数的对数, 因此没有必要研究对数函数在 $-\infty$ 附近的行为!) 同样, 我们来看看 $y=\ln(x)$ 的图像, 如图 9-3 所示.

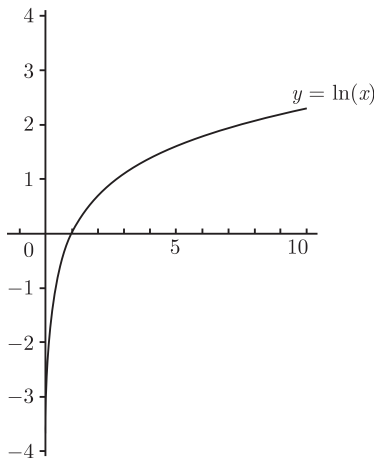


图 9-3

再一次地, 要注意到该曲线绝不会触碰到 y 轴, 尽管看起来它好像是. 它只是非常非常接近 y 轴. 不管怎样, 我们看起来好像有

$$\boxed{\lim_{x\to\infty} \ln(x) = \infty.}$$

这一点实际上很容易直接表明. 你认为 $\ln(x)$ 会达到 1000 吗? 当然会: $\ln\left(\mathrm{e}^{1000}\right) = 1000$. 同样的技巧适用于任意的数 N. 取 $x = \mathrm{e}^N$, 你就会发现 $\ln(x) = \ln\left(\mathrm{e}^N\right) = N$. 因此, $\ln(x)$ 会变多大是没有极限的, 当 $x\to\infty$ 时, 它趋于 $\infty\cdots\cdots$ 但有多快呢?

我们很容易看出其增长速度相当慢. 正如我们刚刚注意到的, $\ln\left(\mathrm{e}^{1000}\right) = 1000$. 数 e^{1000} 是极大的正数 (比宇宙中的原子数目还要大), 而其对数仅为 1000. 简直是把它打回原形!

更确切地说, 事实表明, $\ln(x)$ 趋于无穷大的速度比 x 的**任意正次幂**都要慢很多, 甚至如 $x^{0.0001}$. 因此, 如果你取 $\ln(x)$ 和 x 的任意正次幂的比, 那么该比值应该会很小 (至少当 x 非常大时会很小). 用符号表示, 我们有

对数函数增长缓慢: 不管 a 有多小, $\boxed{\text{如果 } a > 0, \lim_{x\to\infty} \frac{\ln(x)}{x^a} = 0.}$

正如在指数函数中那样, 我们不难将该式扩展成一个更一般的形式:

$$\lim_{x\to\infty} \frac{\text{任何正的多项式型的对数函数}}{\text{具有正的幂次的多项式型}} = 0.$$

这适用于任何底数 $b > 1$ 的对数函数, 而不只是自然对数. (这是因为有换底法则.) 例如,

$$\lim_{x\to\infty} \frac{\log_7(x^3 + 3x - 1)}{x^{0.1} - 99} = 0,$$

尽管 $x^{0.1}$ 非常小.

事实上, 不应该奇怪对数函数增长缓慢, 毕竟我们已经知道指数函数增长迅速, 而对数函数和指数函数互为反函数. 更确切地说, 如果你取 $\ln(x)/x^a$ 并用 e^t 替换 x, 那么会得到

$$\lim_{x\to\infty} \frac{\ln(x)}{x^a} = \lim_{t\to\infty} \frac{\ln(\mathrm{e}^t)}{(\mathrm{e}^t)^a} = \lim_{t\to\infty} \frac{t}{\mathrm{e}^{at}} = 0.$$

最后一个极限是 0, 因为分母中指数函数 e^{at} 的增长比分子中多项式 t 的增长要快很多. 这样我们就证明了, 指数函数增长迅速这一事实会自动导出对数函数增长缓慢这一结论.

9.4.6 对数函数在 0 附近的行为

有人可能想说, $\ln(0) = -\infty$, 但这是不正确的, 因为 $\ln(0)$ 无定义. 另一方面, 前述 $y = \ln(x)$ 的图像暗示了

$$\boxed{\lim_{x\to 0^+} \ln(x) = -\infty.}$$

在这里, 你需要使用右极限, 因为 $\ln(x)$ 在 $x < 0$ 上没有定义. 不过再一次地, 我们还需要再多说一些. 当 $x \to 0^+$ 时, $\ln(x)$ 当然趋于 $-\infty$, 但有多快呢? 例如, 考虑极限

$$\lim_{x\to 0^+} x\ln(x).$$

如果你只是将 0 代入上式, 这根本不起作用, 因为 $\ln(0)$ 不存在. 当 x 是一个比 0 稍大一点的数时, 量 x 很小而 $\ln(x)$ 是一个很大的负数. 当你用一个很大的数和一个很小的数相乘时会怎样呢? 任何情况都可能发生, 取决于那些数有多么小和多么大.

下面是一个求解上述问题的方法. 我们用 $1/t$ 替换 x, 于是当 $x \to 0^+$ 时, 可以看到 $t \to \infty$. 因此, 有

$$\lim_{x\to 0^+} x\ln(x) = \lim_{t\to\infty} \frac{1}{t}\ln\left(\frac{1}{t}\right).$$

当然, $\ln(1/t)$ 正是 $\ln(1) - \ln(t)$, 又由于 $\ln(1) = 0$, 它便等于 $-\ln(t)$. 因此, 得到

$$\lim_{x\to 0^+} x\ln(x) = \lim_{t\to\infty} \frac{1}{t}\ln\left(\frac{1}{t}\right) = \lim_{t\to\infty} \frac{-\ln(t)}{t} = 0.$$

由于对数函数增长缓慢, 故极限是 0.

用 $1/t$ 替换 x 这一技巧可以将对数函数在 0 附近的行为转换为在 ∞ 附近的行为, 因为 $\ln(1/t) = -\ln(t)$. 你可以用它来证明下列原理, 上述例子就是该原理的一个特例:

对数函数在 0 附近 "增长" 缓慢: 不管 a 有多小, $\boxed{\text{如果 } a > 0, \lim_{x\to 0^+} x^a\ln(x) = 0.}$

(我把 "增长" 加上引号, 是因为当 $x \to 0^+$ 时, $\ln(x)$ 实际上是向下增长到 $-\infty$.) 再一次地, 你可以用多项式型来替换 x^a, 只要当 $x \to 0^+$ 时, 它变得非常小, 并且可以用任意的其他底数 $b > 1$ (也就是说, 不限于底数 e) 的 "$\log_b$" 替换 "ln".

9.5 取对数求导法

处理像 $f(x)^{g(x)}$ 这样底数和指数均为 x 的函数的导数问题时, 取对数求导法是一个有用的技巧. 毕竟, 像

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^{\sin(x)})$$

这样的问题用我们之前的方法如何能够解得出来? 根本无从下手. 不过幸好我们还有这些很好的对数法则, 它们能将指数拉下马来. 如果令 $y = x^{\sin(x)}$, 根据 9.1.4 节的对数法则 5, 则

$$\ln(y) = \ln(x^{\sin(x)}) = \sin(x)\ln(x).$$

现在, 对等号两边关于 x 做隐函数求导:

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln(y)) = \frac{\mathrm{d}}{\mathrm{d}x}(\sin(x)\ln(x)).$$



先来看看右边的部分. 它是一个 x 的函数且需要用乘积法则来求解; 你可以验证一下求导结果应该是 $\cos(x)\ln(x) + \sin(x)/x$. 再来看看左边的部分. 为了对 $\ln(y)$ 关于 x (而不是 y) 求导, 我们应该使用链式求导法则. 设 $u = \ln(y)$, 这样 $\mathrm{d}u/\mathrm{d}y = 1/y$. 我们需要求出 $\mathrm{d}u/\mathrm{d}x$; 根据链式求导法则,

$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x}.$$

因此, 对方程 $\ln(y) = \sin(x)\ln(x)$ 进行隐函数求导后得到

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \cos(x)\ln(x) + \frac{\sin(x)}{x}.$$

现在只需要用 y 和等号两边相乘, 然后用 $x^{\sin(x)}$ 替换 y:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \left(\cos(x)\ln(x) + \frac{\sin(x)}{x}\right)y = \left(\cos(x)\ln(x) + \frac{\sin(x)}{x}\right)x^{\sin(x)}.$$

这就是我们要找的答案了. (顺便提一下, 还可以用另外一种方法来解此题. 不是使用变量 y, 直接使用公式 $A = \mathrm{e}^{\ln(A)}$ 来写出

$$x^{\sin(x)} = \mathrm{e}^{\ln(x^{\sin(x)})} = \mathrm{e}^{\sin(x)\ln(x)}.$$



接下去我留给你来使用乘积法则和链式求导法则对右边关于 x 求导. 完成后, 你应该用 $x^{\sin(x)}$ 来替换 $\mathrm{e}^{\sin(x)\ln(x)}$ 并检验你是否得到与前面一样的答案.)



让我们来回顾一下这种技巧吧. 假设要关于 x 求导函数

$$y = f(x)^{g(x)},$$

其中底数 f 和指数 g 都含有变量 x. 以下是你需要做的:

(1) 设 y 是想要求导的 x 的函数. 对等号两边取 (自然) 对数. 右边的指数 g 得以移下来, 这样得到

$$\ln(y) = g(x)\ln(f(x)).$$

(2) 对等号两边关于 x 做隐函数求导. 右边常常会用到乘积法则和 (至少) 链式求导法则. 左边的结果则总是 $(1/y)(\mathrm{d}y/\mathrm{d}x)$. 因此, 你会得到

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \text{关于 } x \text{ 的一堆东西}$$

(3) 用 y 和等式两边相乘会得到单独的 $\mathrm{d}y/\mathrm{d}x$ 这一项, 然后用原始的表达式 $f(x)^{g(x)}$ 替换 y, 你就完成了求解.



下面是另外一个例子:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left((1+x^2)^{1/x^3}\right)$$

是什么呢? 第一步, 设 $y = \left(1+x^2\right)^{1/x^3}$, 然后对等式两边取对数, 这样会使指数移下来, 得到

$$\ln(y) = \ln\left((1+x^2)^{1/x^3}\right) = \frac{1}{x^3}\ln(1+x^2) = \frac{\ln(1+x^2)}{x^3}.$$

第二步是对等式两边关于 x 作隐函数求导. 如往常一样, 左边变为 $(1/y)\,(\mathrm{d}y/\mathrm{d}x)$, 但我们必须对右边使用商法则. 首先, 使用链式求导法则对 $z = \ln\left(1+x^2\right)$ 求导: 如果 $u = 1+x^2$, 那么 $z = \ln(u)$, 故

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{u}(2x) = \frac{2x}{1+x^2}.$$

现在可以使用商法则. 你应该检验一下, 当对上述方程 $\ln(y) = \ln\left(1+x^2\right)/x^3$ 作隐函数求导时, 是否得到了 (经简化之后)

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^3\dfrac{2x}{1+x^2} - 3x^2\ln(1+x^2)}{(x^3)^2} = \frac{2x^2 - 3(1+x^2)\ln(1+x^2)}{x^4(1+x^2)}.$$

最后, 用 y 和等式两边相乘, 并用 $\left(1+x^2\right)^{1/x^3}$ 替换 y, 得到

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{(2x^2 - 3(1+x^2)\ln(1+x^2))y}{x^4(1+x^2)}\\ &= \frac{(2x^2 - 3(1+x^2)\ln(1+x^2))(1+x^2)^{1/x^3}}{x^4(1+x^2)}\\ &= \frac{(2x^2 - 3(1+x^2)\ln(1+x^2))}{x^4(1+x^2)^{1-1/x^3}}.\end{aligned}$$

这样就完成了求解.

即便底数和指数都不是 x 的函数, 取对数求导法仍会非常有用. 如果你的函数非常复杂并涉及很多幂函数 (像 x^2) 和指数函数 (像 e^x) 的乘积和商, 那么取对数求导法就可能会帮得上忙. 例如,



$$\text{如果 } y = \frac{(x^2-3)^{100}3^{\sec(x)}}{2x^5(\log_7(x)+\cot(x))^9}, \text{那么 } \frac{\mathrm{d}y}{\mathrm{d}x} \text{ 是什么?}$$

你一定在想, 我准是在开玩笑. 我们怎么可能对这样糟糕的表达式求导呢? 但使用取对数求导法就可以. 对等式两边取自然对数, 你会发现右边变得容易处理多了 (只要你还记得对数法则):

$$\begin{aligned}\ln(y) &= \ln\left(\frac{(x^2-3)^{100}3^{\sec(x)}}{2x^5(\log_7(x)+\cot(x))^9}\right)\\ &= \ln((x^2-3)^{100}) + \ln(3^{\sec(x)}) - \ln(2) - \ln(x^5) - \ln((\log_7(x)+\cot(x))^9)\\ &= 100\ln(x^2-3) + \sec(x)\ln(3) - \ln(2) - 5\ln(x) - 9\ln(\log_7(x)+\cot(x)).\end{aligned}$$

在继续阅读之前, 确保你理解了这些对数的操作. 不管怎样, 现在我们可以对该表达式关于 x 作隐函数求导了:

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln(y)) = \frac{\mathrm{d}}{\mathrm{d}x}(100\ln(x^2-3)+\sec(x)\ln(3) \\ -\ln(2)-5\ln(x)-9\ln(\log_7(x)+\cot(x))).$$

如往常一样, 左边是 $(1/y)\,(\mathrm{d}y/\mathrm{d}x)$, 右边则让我们来逐项地看.

- 第一项是 $100\ln\left(x^2-3\right)$. 这正是一个简单的链式求导法则的练习, 容易看出其导数就是 $100\times 2x/(x^2-3)$, 也就是 $200x/\left(x^2-3\right)$.
- 第二项是 $\sec(x)\ln(3)$. 在准备使用乘积法则之前, 要注意到 $\ln(3)$ 是一个常数, 因此你实际上可以只求 $\sec(x)$ 的导数, 然后再和 $\ln(3)$ 相乘得到 $\ln(3)\sec(x)\tan(x)$.
- 第三项是 $-\ln(2)$. 它是一个常数, 故其导数就是 0.
- 第四项是 $-5\ln(x)$. 其导数为 $-5/x$.
- 第五项是 $-9\ln(\log_7(x)+\cot(x))$. 我称之为 z, 我们需要使用链式求导法则. 尽管你应该能够自己求出, 但我还是将细节列出来. 令 $u=\log_7(x)+\cot(x)$, 故 $z=-9\ln(u)$. 那么我们有

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = -\frac{9}{u}\left(\frac{1}{x\ln(7)}-\csc^2(x)\right) \\ = \frac{9}{\log_7(x)+\cot(x)}\left(\csc^2(x)-\frac{1}{x\ln(7)}\right).$$

综合起来会得到

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{200x}{x^2-3}+\ln(3)\sec(x)\tan(x)-\frac{5}{x} \\ +\frac{9}{\log_7(x)+\cot(x)}\left(\csc^2(x)-\frac{1}{x\ln(7)}\right).$$

现在, 用 y 与之相乘会得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{200x}{x^2-3}+\ln(3)\sec(x)\tan(x)-\frac{5}{x} \right.\\ \left. +\frac{9}{\log_7(x)+\cot(x)}\left(\csc^2(x)-\frac{1}{x\ln(7)}\right)\right)\times y.$$

最后, 用原始的 (可怕的) 表达式替换 y 会得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{200x}{x^2-3}+\ln(3)\sec(x)\tan(x)-\frac{5}{x} \right.\\ \left. +\frac{9}{\log_7(x)+\cot(x)}\left(\csc^2(x)-\frac{1}{x\ln(7)}\right)\right)\times\frac{(x^2-3)^{100}3^{\sec(x)}}{2x^5(\log_7(x)+\cot(x))^9}.$$

这看上去有点难以处理, 但想象一下要是不用取对数求导法的话 ……

x^a 的导数

现在我们终于可以来证明之前一直不作证明就接受的一件事情了：对于任意的数 a (而不只是我们之前见到整数), 有

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}(x^a) = ax^{a-1}.}$$

不妨假设 $x > 0$. 现在使用取对数求导法：设 $y = x^a$, 这样 $\ln(y) = a\ln(x)$. 如果你对两边作隐函数求导, 会得到

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{a}{x}.$$

现在用 y 与两边相乘并用 x^a 替换 y：

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{ay}{x} = \frac{ax^a}{x} = ax^{a-1}.$$

这就是我们想要的, 至少当 $x > 0$ 时. 当 $x \leqslant 0$ 时, 我们有一点小问题. 例如, 你不能取 $(-1)^{1/2}$, 因为这是一个负数的平方根. $(-1)^{\sqrt{2}}$ 究竟又会是什么呢？事实上, 如果不使用复数 (毕竟, 直到第 28 章我们才会学到), 只有当 a 是一个带有一个奇分母 (在消去公因子之后) 的有理数时, 对于 $x < 0$, x^a 才说得通. 例如, 对于 $x < 0$, $x^{5/3}$ 说得通, 因为你总是可以求一个立方根 —— 由于 3 是奇数, 所以这没有问题. 在 $x < 0$ 而 x^a 说得通的情况下, 结果表明, 它或者是 x 的偶函数或者是 x 的奇函数, 而你可以使用这个事实证明其导数仍是 ax^{a-1}.

这里有一些使用以上公式的简单例子. 如果定义域是 $(0, \infty)$, 那么 $x^{\sqrt{2}}$ 关于 x 的导数是什么呢？x^{π} 呢？使用公式可知, 对于 $x > 0$,



$$\frac{\mathrm{d}}{\mathrm{d}x}(x^{\sqrt{2}}) = \sqrt{2}x^{\sqrt{2}-1} \quad 和 \quad \frac{\mathrm{d}}{\mathrm{d}x}(x^{\pi}) = \pi x^{\pi-1}.$$

这和我们之前所做的没什么实质性区别, 只是我们现在可以处理非整数的指数了.

9.6 指数增长和指数衰变

我们已经看到, 连续计复利的银行账户的增长是指数式的. 不过, 指数增长不只限于人类社会, 它也见于自然界. 例如, 在一定条件下, 动物种群总数, 如兔子 (和人类) 会呈指数增长. 此外, 还有指数衰变, 其中一个量以指数方式变得越来越小 (我们很快就能看到这是什么意思). 放射性衰变就是如此, 这使得科学家能够确定古代器物、化石或岩石的年龄.

以下是基本思想. 假设 $y = \mathrm{e}^{kx}$. 那么, 正如我们在 9.3.1 节的开头部分看到的, $\mathrm{d}y/\mathrm{d}x = k\mathrm{e}^{kx}$. 我们可以将等式右边写作 ky, 因为 $y = \mathrm{e}^{kx}$. 也就是说,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = ky.$$

这是**微分方程**的一个例子. 也就是说, 它是一个涉及导数的方程. 我们将在第 30 章看到更多的微分方程, 但现在, 暂且让我们把精力集中在这一个上面. 还有其他什么函数也满足上述方程呢? 我们知道 $y = \mathrm{e}^{kx}$ 满足, 但一定还有其他函数也满足. 例如, 如果 $y = 2\mathrm{e}^{kx}$, 那么 $\mathrm{d}y/\mathrm{d}x = 2k\mathrm{e}^{kx}$, 其结果再次为 ky. 更一般地来说, 如果 $y = A\mathrm{e}^{kx}$, 那么 $\mathrm{d}y/\mathrm{d}x = Ak\mathrm{e}^{kx}$, 也就是 ky. 事实表明, 这是可以得到 $\mathrm{d}y/\mathrm{d}x = ky$ 的**唯一**途径:

如果 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = ky$, 那么 $y = A\mathrm{e}^{kx}$, 其中 A 为某个常数.

我们将在 30.2 节说明为什么会这样. 同时, 让我们再来更深入地看看微分方程 $\mathrm{d}y/\mathrm{d}x = ky$. 首先要做的是将变量 x 变为 t, 这样可以看到

$$\frac{\mathrm{d}y}{\mathrm{d}t} = ky.$$

这意味着, y 的变化率等于 ky. 这太有趣了! 一个量变化的速率取决于这个量的大小. 如果这个量越大, 那么它就会增长得越快 (假设 $k > 0$). 在动物种群总数增长的情况中, 这是说得通的: 兔子越多, 它们就可以繁殖得越多. 如果你有两倍的兔子, 那么在任意给定的时间周期中, 它们也会繁殖出两倍的兔子. 数 k 被称为**增长常数**, 它控制着兔子繁殖得多快. 它们性致越高, k 越大!

9.6.1　指数增长

假设有一个种群以指数增长. 用符号表示, 设 P (或 $P(t)$, 如果你喜欢) 是在时刻 t 时的总数, 并设 k 是增长常数. P 的微分方程为

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP.$$

这和前面方框中的微分方程是一样的, 除了一些符号有所不同. 这里我们有 P 而不是 y; 有 t 而不是 x. 不过不要紧, 我们向来善于随机应变; 我们只需在解 $y = A\mathrm{e}^{kx}$ 中做同样的替换即可. 这样对于某个常数 A, $P = A\mathrm{e}^{kt}$. 现在, 当 $t = 0$ 时, 我们有 $P = A\mathrm{e}^{k(0)} = A\mathrm{e}^0 = A$, 因为 $\mathrm{e}^0 = 1$. 这意味着, A 是初始的总数, 即在时刻 0 时的总数. 习惯上, 我们会用新的符号表示这个变量, 用 P_0 来代替 A, 以表明它代表的是在时刻 0 时的总数. 综合在一起, 我们有

指数增长方程: $P(t) = P_0\mathrm{e}^{kt}$.

其中 P_0 是初始的总数, k 是增长常数.

此公式很容易应用到实际中, 只要你知道指数法则和对数法则 (参见 9.1.1 节和 9.1.4 节). 例如, 如果你知道三年前兔子的总数是 1000 只, 而现在增长至 64 000



只, 那么从现在算起, 一年之后兔子总数会是多少呢? 此外, 总数从 1000 增长至 400 000 需要多长时间呢?

好吧, 我们有 $P_0 = 1000$, 因为这是初始的总数. 故上述方框中的方程变为 $P(t) = 1000\mathrm{e}^{kt}$. 但问题是, 我们不知道 k 是什么. 我们知道的是, 当 $t = 3$ 时, $P = 64\,000$, 因此将其代入:

$$64\,000 = 1000\mathrm{e}^{3k}.$$

这意味着, $\mathrm{e}^{3k} = 64$. 对两边取对数可得 $3k = \ln(64)$, 故 $k = \frac{1}{3}\ln(64)$. 事实上, 如果你写出 $\ln(64) = \ln\left(2^6\right) = 6\ln(2)$, 那么就可以将其化简为 $k = 2\ln(2)$. 这意味着, 对于任意的时刻 t,

$$P(t) = 1000\mathrm{e}^{2\ln(2)t}.$$

现在, 可以求解该问题了. 对于第一部分, 我们想要知道从现在开始的一年中会发生什么情况. 这实际上是从初始时刻开始的第四年, 故设 $t = 4$, 得到

$$P(4) = 1000\mathrm{e}^{2\ln(2)\times 4} = 1000\mathrm{e}^{8\ln(2)}.$$

这里有一个小窍门: 将 $8\ln(2)$ 写作 $\ln\left(2^8\right) = \ln(256)$, 故

$$P(4) = 1000\mathrm{e}^{\ln(256)} = 1000 \times 256 = 256\,000.$$

这里使用了重要公式: 对于任意的数 $A > 0$, $\mathrm{e}^{\ln(A)} = A$. 因此结论是, 从现在算起一年后, 兔子总数将变为 256 000. 现在, 我们来处理问题的第二部分. 我们想要知道需要多长时间总数会增至 400 000, 故设 $P = 400\,000$, 得到

$$400\,000 = 1000\mathrm{e}^{2\ln(2)t}.$$

这变为 $\mathrm{e}^{2\ln(2)t} = 400$. 为了求解这一方程, 我们对等式两边取对数, 得到 $2\ln(2)\,t = \ln(400)$, 这意味着

$$t = \frac{\ln(400)}{2\ln(2)}.$$

这就是总数从 1000 增至 400 000 所需的年数, 但这并不是很直观. 你可以使用计算器算出一个近似值. 但如果你手边没有的话, 就需要知道 $\ln(5)$ 近似为 1.6, 而 $\ln(2)$ 近似为 0.7. 我们先写出 $400 = 20^2$, 这样 $\ln(400) = \ln\left(20^2\right) = 2\ln(20)$. 不过, 我们还可以做得更好, $\ln(20) = \ln(4\times 5) = \ln(4) + \ln(5) = 2\ln(2) + \ln(5)$. 综上所述, 得到

$$t = \frac{\ln(400)}{2\ln(2)} = \frac{2(2\ln(2) + \ln(5))}{2\ln(2)} = 2 + \frac{\ln(5)}{\ln(2)}.$$

使用近似值, 得到

$$t \approx 2 + \frac{1.6}{0.7} = 2 + \frac{16}{7} = 4\frac{2}{7}.$$

因此, 尽管需要四年兔子总数才能达到 256 000, 但只需要再有大约七分之二年 (也就是三个半月左右) 就能达到 400 000. 这就是指数增长的威力 ……

9.6.2　指数衰变

现在让我们掉过头来看看指数衰变. 作为铺垫, 我要告诉你, 有些元素的原子具有放射性. 它们像微小的定时炸弹：一段时间后, 原子核分裂, 它们变成别的元素, 同时释放出能量. 唯一的问题是, 你无法知道原子核何时会分裂 (下面我们将不再说 “分裂”, 而称之为 “衰变”). 你知道的只是经过给定的一段时间, 存在一定的概率会发生衰变.

例如, 你有一种特定元素的原子, 它在任意的七年周期内衰变的概率是 50%. 因此, 如果在一个盒子里你有这一个原子, 关上盒子, 并在七年后打开, 那么它已经衰变的概率就是五五开. 当然, 看到一个单独的原子相当难! 因此, 我们假设, 更现实一点, 你有一万亿个原子 (顺便一提, 这仍然只是微乎其微的一小点原料). 你将它们放入盒子, 七年后回来. 你期待发现什么？好吧, 大概有一半的原子应该已经衰变了, 而另一半仍是完好的. 因此, 你应该有大概一万亿的一半的原始原子. 再过七年, 你再来看时又会怎样呢？剩余的一半原子会仍然完好, 即留给你的是一万亿的四分之一的原始原子. 每隔七年, 你失去所剩样本的一半的原子.

因此, 让我们试着写出一个方程来对该问题建模. 如果 $P(t)$ 是原子在时刻 t 时的数量 (总数？), 那么我可以断言,

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -kP,$$

其中 k 是某个常数. 这说的是, P 的变化率是 P 的负倍数. 也就是说, P 是以一个和 P 成比例的速率衰变的. 你拥有的原子数量越多, 它们衰变得就越快. 这和上述例子是一致的：在第一个七年中, 我们失去了一万亿原子中的一半, 而在下一个七年中, 我们只失去了一万亿的四分之一的原子, 再过七年, 我们只会失去一万亿的八分之一的原子. 我们拥有的越多, 失去的也就越多. 不管怎样, 上述微分问题的解是

$$P(t) = P_0\mathrm{e}^{-kt},$$

其中 P_0 是原子的原始数量 (在 $t = 0$ 时). 这和上一节中指数增长的方程是一样的, 除了我们将增长常数 k 代之以一个负的常数 $-k$, 称为**衰变常数**.

在上例中, 我们知道对于这种原子的任何样本, 需要七年时间数量才会减半. 这个时间长度被称为原子 (或原料) 的**半衰期**. 在上述方程中, 这意味着, 如果你开始时有 P_0 个原子, 那么七年后, 你会剩下 $\frac{1}{2}P_0$ 个原子. 因此, 设 $t = 7$ 且上式中 $P(7) = \frac{1}{2}P_0$, 我们有

$$\frac{1}{2}P_0 = P_0\mathrm{e}^{-k(7)}.$$

现在, 从等号两边消去因子 P_0 并对两边取对数, 我们得到

$$\ln\left(\frac{1}{2}\right) = -7k.$$

由于 $\ln(1/2) = \ln(1) - \ln(2) = -\ln(2)$, 上述方程变为

$$k = \frac{\ln(2)}{7}.$$

这意味着, 在这种情况下,

$$P(t) = P_0\mathrm{e}^{-t(\ln(2)/7)}.$$

现在, 我们将以上情况一般化. 假设有另一种放射性原料, 它的半衰期是 $t_{1/2}$ 年. 这意味着, 任何大小的原料样本的一半会在 $t_{1/2}$ 年后衰变. 但这并不意味着整个样本会在两倍的那么多年后全部衰变掉! 不管怎样, 依据和上一段中相同的推理, 我们可以证明 $k = \ln(2)/t_{1/2}$. 总之,

> 对于半衰期为 $t_{1/2}$ 的放射性衰变, $P(t) = P_0\mathrm{e}^{-kt}$, 其中 $k = \dfrac{\ln(2)}{t_{1/2}}$.

例如, 如果原料的半衰期仍是七年, 开始时有 50 磅原料, 十年后还剩多少呢? 需要多久, 原料会减少为 1 磅呢? 我们知道 $t_{1/2} = 7$, 故 $k = \ln(2)/7$, 正如我们之前所见. 由于 $P_0 = 50$ (单位为磅), 衰变方程 $P(t) = P_0\mathrm{e}^{-kt}$ 变为



$$P(t) = 50\mathrm{e}^{-t(\ln(2)/7)}.$$

故当 $t = 10$ 时, 有

$$P(10) = 50\mathrm{e}^{-10\ln(2)/7}.$$

也就是说, 原料缩减为 $50\mathrm{e}^{-10\ln(2)/7}$ 磅. 如果代入近似值 $\ln(2) \approx 0.7$, 那么可以看到还剩大约 $50\mathrm{e}^{-1}$ 磅, 这可以进而近似为大概 18.4 磅.

至于问题的第二部分, 现在要求出需要多久原料会缩减为 1 磅, 故在上面 $P(t)$ 的方程中设 $P(t) = 1$, 得到

$$1 = 50\mathrm{e}^{-t(\ln(2)/7)}.$$

两边同除以 50 并取对数, 得

$$\ln\left(\frac{1}{50}\right) = -\frac{t\ln(2)}{7}.$$

由于 $\ln(1/50) = -\ln(50)$, 有 $-7\ln(50) = -t\ln(2)$. 也就是说,

$$t = \frac{7\ln(50)}{\ln(2)}.$$

可以使用之前的近似值 $\ln(5) \approx 1.6$ 和 $\ln(2) \approx 0.7$ 对此进行估算. 我们写出 $\ln(50) =$

$\ln(2\times5\times5)=\ln(2)+2\ln(5)$, 从而得到

$$t=\frac{7\ln(50)}{\ln(2)}=\frac{7(\ln(2)+2\ln(5))}{\ln(2)}=7+\frac{14\ln(5)}{\ln(2)}$$
$$\approx 7+\frac{14(1.6)}{0.7},$$

其结果是 39 年. 因此, 样本大概需要 39 年从 50 磅衰变到 1 磅. 顺便说一下, 39 年比 $5\frac{1}{2}$ 个半衰期略微多一点 (因为一个半衰期是七年). 因此, 如果有 50 磅另一种不同的原料, 其半衰期为十年, 那么该原料将需要比 55 年略微多一点的时间衰变到 1 磅. (实际的结果是 $10\ln(50)/\ln(2)$ 年, 这约为 $56\frac{1}{2}$ 年.)

9.7 双曲函数

现在让我们改变一下行径, 来探讨一下所谓的**双曲函数**. 它们实际上是伪装的指数函数, 但它们在很多方面又和三角函数非常相似. 我们不会用到太多的双曲函数, 但它们偶尔会出现, 因此最好还是熟悉一下它们.

我们先来定义双曲余弦函数和双曲正弦函数：

$$\boxed{\cosh(x)=\frac{\mathrm{e}^x+\mathrm{e}^{-x}}{2}} \quad 和 \quad \boxed{\sinh(x)=\frac{\mathrm{e}^x-\mathrm{e}^{-x}}{2}}$$

完全不需要三角形! 毕竟, 这根本不是三角学.[①] 这些函数的行为有些像普通的函数, 但又不完全是. 例如, 如果平方 $\cosh(x)$ 和 $\sinh(x)$, 你会发现

$$\cosh^2(x)=\left(\frac{\mathrm{e}^x+\mathrm{e}^{-x}}{2}\right)^2=\frac{\mathrm{e}^{2x}+\mathrm{e}^{-2x}+2}{4},$$

和

$$\sinh^2(x)=\left(\frac{\mathrm{e}^x-\mathrm{e}^{-x}}{2}\right)^2=\frac{\mathrm{e}^{2x}+\mathrm{e}^{-2x}-2}{4}.$$

(我们使用了 $\mathrm{e}^x\mathrm{e}^{-x}=1$ 这个事实.) 然后取这两个量的差：

$$\cosh^2(x)-\sinh^2(x)=\frac{\mathrm{e}^{2x}+\mathrm{e}^{-2x}+2}{4}-\frac{\mathrm{e}^{2x}+\mathrm{e}^{-2x}-2}{4}=\frac{4}{4}=1.$$

这样就证明了, 对于任意的 x,

$$\boxed{\cosh^2(x)-\sinh^2(x)=1.}$$

这和原来的三角恒等式不太一样 —— 减号让一切变得不同. (确实, $x^2-y^2=1$ 是一个双曲方程.)

① 事实上, 有一个几何学分支, 称为**双曲几何学**, 其中三角形有些古怪的性质, 它们引出了双曲函数.

其微积分性质又如何呢？让我们来对 $y = \sinh(x)$ 求导. 我们需要用到 e^{-x} 的导数是 $-\mathrm{e}^{-x}$ 的事实：

$$\frac{\mathrm{d}}{\mathrm{d}x}\sinh(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2}\right) = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2} = \cosh(x).$$

因此, 双曲正弦函数的导数就是双曲余弦函数. 这就好像原来常规正弦函数和余弦函数的情况. 另一方面,

$$\frac{\mathrm{d}}{\mathrm{d}x}\cosh(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}\right) = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2} = \sinh(x).$$

要是这里是普通的三角函数的话, 那么其导数将是负的双曲正弦函数, 但实际上我们在这里没有负号. 不管怎样, 我们证明了

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\sinh(x) = \cosh(x)} \quad 和 \quad \boxed{\frac{\mathrm{d}}{\mathrm{d}x}\cosh(x) = \sinh(x).}$$

现在, 来看看这些函数的图像吧. 首先, 你应该试着让自己相信, $\cosh(x)$ 是 x 的偶函数, 而 $y = \sinh(x)$ 是 x 的奇函数. (只需将 $-x$ 代入看看会发生什么就一目了然了.) 此外, $\cosh(0) = 1$ 且 $\sinh(0) = 0$ (也请检验). 最后, 我们注意到



$$\lim_{x\to\infty}\cosh(x) = \lim_{x\to\infty}\frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}.$$

其中 e^x 趋于 ∞, 而 e^{-x} 趋于 0. 整体的效果就是极限是 ∞. 同理适用于 $\sinh(x)$, 因此它们的图像看起来如图 9-4 所示.

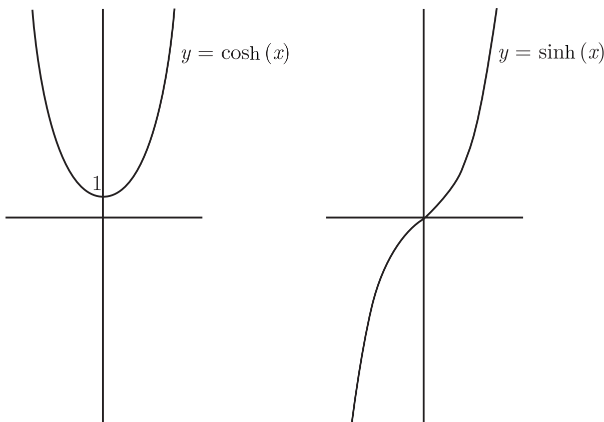


图 9-4

当然, 你可以用 $\sinh(x) / \cosh(x)$ 来定义 $\tanh(x)$, 还有作为各种倒数的 $\mathrm{sech}(x)$、$\mathrm{csch}(x)$ 及 $\coth(x)$. 我们可以通过适当地替换指数函数来区分双曲正割函数、双曲余割函数和双曲余切函数. 例如,

$$\text{sech}(x) = \frac{1}{\cosh(x)} = \frac{1}{\dfrac{e^x + e^{-x}}{2}} = \frac{2}{e^x + e^{-x}},$$

你可以使用链式求导法则或乘积法则对它求导. 我们还有联系这些函数的恒等式, 其中最重要的一个就是

$$1 - \tanh^2(x) = \text{sech}^2(x).$$

这可以直接从恒等式 $\cosh^2(x) - \sinh^2(x) = 1$ 中推出：两边同除以 $\cosh^2(x)$. 现在, 我要列出其他双曲函数的导数并展示它们的图像了 —— 我留给你去检验所有的导数是正确的以及至少一个图像是说得通的. 首先是导数：

$$\boxed{\frac{d}{dx}\tanh(x) = \text{sech}^2(x)} \quad \boxed{\frac{d}{dx}\text{sech}(x) = -\,\text{sech}(x)\tanh(x)}$$

$$\boxed{\frac{d}{dx}\text{csch}(x) = -\,\text{csch}(x)\coth(x)} \quad \boxed{\frac{d}{dx}\coth(x) = -\,\text{csch}^2(x).}$$

然后是图 9-5.

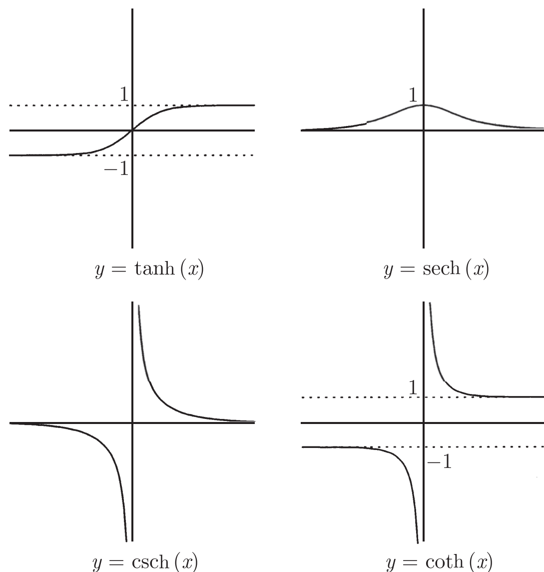


图　9-5

由函数的定义可以看到, 除了双曲余弦函数和双曲正割函数是偶函数外, 所有的双曲三角函数都是奇函数. 这和原来常规的三角函数的情况相同! 此外, $y = \tanh(x)$ 和 $y = \coth(x)$ 都在 $y = 1$ 和 $y = -1$ 处有水平渐近线, 而 $y = \text{sech}(x)$ 和 $y = \text{csch}(x)$ 在 $y = 0$ 处都有一条水平渐近线.

第 10 章　反函数和反三角函数

在上一章中, 我们研究了指数函数和对数函数, 并围绕以下事实讨论了很多, 那就是 e^x 和 $\ln(x)$ 互为反函数. 在本章, 我们将先看一下反函数更一般的性质, 然后再详细讨论反三角函数 (及反双曲函数). 以下就是我们的计划：

- 使用导数证明一个函数有反函数;
- 求反函数的导数;
- 逐个来看反三角函数;
- 反双曲函数.

10.1　导数和反函数

在 1.2 节中, 我们回顾了反函数的基本知识. 我强烈建议你在继续阅读之前再快速地浏览一下那节, 让自己重温一下大致思想. 现在我们已经了解了一些微积分知识, 对此就有更多可说的了. 特别是, 我们将要讨论导数和反函数之间的两个联系.

10.1.1　使用导数证明反函数存在

假设有一个可导函数 f, 它的导数总是正的. 你认为该函数的图像会是什么样的呢？好吧, 切线的斜率必定处处为正, 故该函数不可能上下起伏：当我们从左向右看时, 它必须是向上的. 换句话说, 该函数一定是**递增的**.

我们会在下一章中 (参见 11.3.1 节及 11.2 节) 证明这个事实, 但现在至少看上去它应该是成立的. 不管怎样, 如果函数 f 总是递增的, 那么它一定满足水平线检验. 没有水平线会与 $y = f(x)$ 相交两次. 由于 f 满足水平线检验, 所以我们知道 f 有反函数. 这就给我们提供了一个证明一个函数有反函数的很好策略：证明它的导数在其定义域上总为正.



例如, 假设在其定义域 $\mathbb{R}$ (整个实轴) 上,

$$f(x) = \frac{1}{3}x^3 - x^2 + 5x - 11.$$

f 有反函数吗？试图在方程 $y = \frac{1}{3}x^3 - x^2 + 5x - 11$ 中调换 x 和 y, 然后求解 y, 无疑会弄得一团糟. (你试着做做看!) 证明 f 有反函数的一个更好方法就是求其导数. 我们得到

$$f'(x) = x^2 - 2x + 5.$$

但那又怎样呢? 好吧, f' 是一个二次函数. 它的判别式为 -16, 这是负的, 因此方程 $f'(x)=0$ 无解. (关于判别式, 参见 1.6 节.) 这意味着, $f'(x)$ 一定总为正或总为负: 它的图像不可能与 x 轴相交. 那好, 它究竟是正的还是负的呢? 由于 $f'(0)=5$, 它一定为正.[①] 也就是说, 对于所有的 x, $f'(x)>0$. 这意味着, f 是递增的. 特别是, f 满足水平线检验, 因此它有反函数.

我们已经看到, 如果对于所有的在其定义域中的 x, $f'(x)>0$, 那么 f 有反函数. 这里还有一些变体. 例如, 如果对于所有的 x, $f'(x)<0$, 那么 $y=f(x)$ 的图像是递减的. 尽管如此, 水平线检验仍然适用: 图像是一直向下的, 所以它不可能掉头向上并与同样的水平线相交两次. 另一个变体是, 其导数在某个位置可能是 0, 但在其他地方都是正的. 这没有问题, 只要其导数不在 0 上逗留太久. 以下就是我们对情况的总结.

导数和反函数: 如果 f 在其定义域 (a,b) 上可导且满足以下条件中的任意一条:

(1) 对于所有的在 (a,b) 中的 x, $f'(x)>0$;

(2) 对于所有的在 (a,b) 中的 x, $f'(x)<0$;

(3) 对于所有的在 (a,b) 中的 x, $f'(x)\geqslant 0$ 且对于有限个数的 x, $f'(x)=0$;

(4) 对于所有的在 (a,b) 中的 x, $f'(x)\leqslant 0$ 且对于有限个数的 x, $f'(x)=0$,

则 f 有反函数. 如果其定义域是 $[a,b]$、$[a,b)$ 或 $(a,b]$ 的形式, 且 f 在整个定义域上连续, 那么如果 f 满足上述四个条件中的任意一条, 它仍然有反函数.

下面是另一个例子. 假设在定义域 $(0,\ \pi)$ 上 $g(x)=\cos(x)$. g 有反函数吗? 首先, $g'(x)=-\sin(x)$. 我们知道, 在区间 $(0,\pi)$ 上, $\sin(x)>0$—— 如果你不相信的话, 只需看一下它的图像. 由于 $g'(x)=-\sin(x)$, 我们看到, 对于所有的在 $(0,\pi)$ 中的 x, $g'(x)<0$. 这意味着, g 有反函数. 事实上, 我们知道在整个的 $[0,\pi]$ 上 g 有反函数, 因为 g 在那里是连续的. 这里的基本思路是, $g(0)=1$, 故 g 始于高度 1; 又由于当 $0<x<\pi$ 时, $g'(x)<0$, 我们知道 g 会立即变得低于 1. 又由于 $g(\pi)=-1$, $g(x)$ 的值会下降至 -1, 并且在这个过程中不会两次到达同一个值. 因此在整个 $[0,\pi]$ 上 g 有反函数. 我们将在 10.2.2 节再次讨论这个函数.

最后一个例子, 在整个 $\mathbb{R}$ 上令 $h(x)=x^3$. 我们知道 $h'(x)=3x^2$, 它不可能是负的. 因此, 对于所有的 x, $h'(x)\geqslant 0$. 幸运的是, 仅当 $x=0$ 时 $h'(x)=0$, 故只有一点使得 $h'(x)=0$. 这就没问题了, 因此 h 仍然有反函数; 事实上, $h^{-1}(x)=\sqrt[3]{x}$.

10.1.2　导数和反函数: 可能出现的问题

我们注意到, 函数的导数可以偶尔是 0, 而该函数仍然有反函数. 但为什么不能允许稍微多一点 $f'(x)=0$ 呢? 例如, 假设 f 定义如下:

① 另一个证明方法是配方: $x^2-2x+5=(x-1)^2+4>0$, 因为所有平方 (如 $(x-1)^2$) 都是非负的.

$$f(x)=\begin{cases}-x^2+1 & \text{如果 } x<0,\\ 1 & \text{如果 } 0\leqslant x<1,\\ x^2-2x+2 & \text{如果 } x\geqslant 1.\end{cases}$$

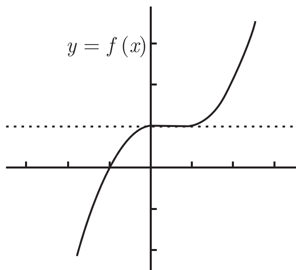


图 10-1

当 $x<0$ 时, 我们有 $f'(x)=-2x$, 它是正的 (因为 x 是负的!). 当 $0<x<1$ 时, 我们有 $f'(x)=0$; 当 $x>1$ 时, 我们可以看到 $f'(x)=2x-2=2(x-1)$, 这一定是正的. 此外, 函数值和导数值在连接点 $x=0$ 和 $x=1$ 处是一致的, 这样就证明了 f 可导且对于所有的 x, $f'(x)\geqslant 0$. (为什么这是可以的, 参见 6.6 节.) 但不幸的是, 它没有通过水平线检验, 故不存在反函数! 我们来检验一下其图像, 如图 10-1 所示. 水平线 y=1 和该图像有无数次相交 (在 x=0 和 x=1 之间且包括这两点的每一点上都相交). 函数 f 在 $[0,1]$ 上是常数, 这与对于那些 x, $f'(x)=0$ 的事实是一致的.

这里还有另一个潜在的问题. 10.1.1 节中的那四个条件都要求定义域是一个形如 (a,b) 的区间. 如果定义域不在一起会怎样呢？不幸的是, 要是那样的话, 结论就全都不成立了. 例如, 如果 $f(x)=\tan(x)$, 那么 $f'(x)=\sec^2(x)$, 这不可能是负的; 然而, 从图像中可以看到 $y=\tan(x)$ 不满足水平线检验. ($y=\tan(x)$ 的图像参见 10.2.3 节.) 因此一般来说, 当函数有不连续点或垂直渐近线时, 上节所述方法就不适用了.

10.1.3 求反函数的导数

如果知道函数 f 有反函数, 我们通常称之为 f^{-1}, 那么该反函数的导数是什么呢？下面就介绍如何求解. 从方程 $y=f^{-1}(x)$ 开始. 你可以将它重新写作 $f(y)=x$. 现在对方程两边关于 x 作隐函数求导得到

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(y))=\frac{\mathrm{d}}{\mathrm{d}x}(x).$$

等号右边很容易求解, 它就是 1. 为了求解左边, 我们使用隐函数求导 (参见第 8 章). 如果设 $u=f(y)$, 那么根据链式求导法则 (注意到 $\mathrm{d}u/\mathrm{d}y=f'(y)$), 我们有

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(y))=\frac{\mathrm{d}}{\mathrm{d}x}(u)=\frac{\mathrm{d}u}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x}=f'(y)\frac{\mathrm{d}y}{\mathrm{d}x}.$$

现在, 等式两边同除以 $f'(y)$, 得到以下定理：

$$\boxed{\text{如果 } y=f^{-1}(x), \text{ 则 } \frac{\mathrm{d}y}{\mathrm{d}x}=\frac{1}{f'(y)}.}$$

如果想要用 x 来表达所有项, 那么必须用 $f^{-1}(x)$ 替换 y, 得到

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}(f^{-1}(x)) = \frac{1}{f'(f^{-1}(x))}.}$$

这意味着, 反函数的导数基本上就是原函数的导数的倒数, 只是对于后面这个导数你必须用 $f^{-1}(x)$ 而不是 x 进行计算.

例如, 设 $f(x) = \frac{1}{3}x^3 - x^2 + 5x - 11$. 在 10.1.1 节中我们已经看到, f 在其定义域 $\mathbb{R}$ 上有反函数. 如果设 $y = f^{-1}(x)$, 那么 $\mathrm{d}y/\mathrm{d}x$ 的一般形式是什么呢? 当 $x = -11$ 时它的值又是什么呢? 为了求解第一部分, 你所要做的只是看到 $f'(x) = x^2 - 2x + 5$, 故

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{f'(y)} = \frac{1}{y^2 - 2y + 5}.$$

注意到, 在这里, 重要的是要用 y 替换 x. 不管怎样, 现在我们可以来求解第二部分了. 我们知道 $x = -11$, 但 y 是什么呢? 由于 $y = f^{-1}(x)$, 我们知道有 $f(y) = x$. 根据 f 的定义, 有

$$\frac{1}{3}y^3 - y^2 + 5y - 11 = -11.$$

现在很明显 $y = 0$ 是该方程的一个解, 并且它**一定**是唯一解, 因为反函数存在. 因此, 当 $x = -11$ 时, 有 $y = 0$, 且

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{y^2 - 2y + 5} = \frac{1}{(0)^2 - 2(0) + 5} = \frac{1}{5}.$$

更正式地, 可以写成 $\left(f^{-1}\right)'(-11) = 1/5$.

现在, 假设 $h(x) = x^3$, 如 10.1.1 节所述. 我们已经在那里看到 h 有反函数, 甚至能把它写出来: $h^{-1}(x) = x^{1/3}$. 当然, 可以直接对 x^a 关于 x 求导, 但还是让我们来试一下上述方法吧. 我们知道 $h'(x) = 3x^2$; 如果 $y = h^{-1}(x)$, 那么

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{h'(y)} = \frac{1}{3y^2}.$$

现在, 可以通过解方程 $x = y^3$ 来求 y, 得到 $y = x^{1/3}$, 并将其代入上述方程得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{3(x^{1/3})^2} = \frac{1}{3x^{2/3}}.$$

这样做是相当愚蠢的, 因为我们可以直接对 $y = x^{1/3}$ 求导, 不用那么麻烦就得到相同的答案. 然而, 知道这种方法能奏效还是蛮不错的.

在继续看另一个例子之前, 我们需要注意到一点, 当 $x = 0$ 时, 反函数的导数不存在, 因为分母 $3x^{2/3}$ 变为零. 因此, 尽管原函数处处可导, 其反函数不一定处处可导: 当 $x = 0$ 时, 其导数不存在. 这对一般而言也是成立的, 而不只是对上述函

数 h. 如果你有一个函数, 它有反函数, 并且原函数在点 (x, y) 处的斜率为 0, 则其反函数在点 (y, x) 处的斜率将会是无限的, 如图 10-2 所示.

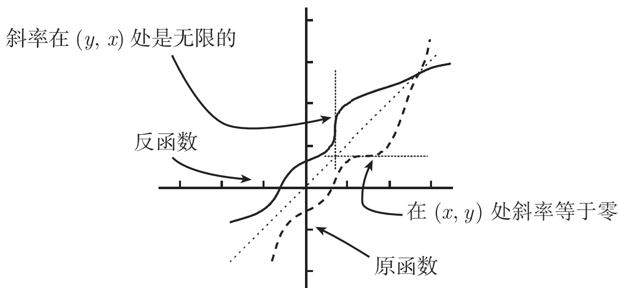


图　10-2

有时候虽然你对一个函数了解不多, 但仍然可以得出有关其反函数的导数的一些信息. 例如, 假设你知道对于一些可逆函数 f, 有 $g(x) = \sin(f^{-1}(x))$, 而你仅知道 f 有 $f(\pi) = 2$ 及 $f'(\pi) = 5$. 但这些信息实际上足够让你求出 $g(2)$ 和 $g'(2)$ 的值了. 具体说, 由于 $f(\pi) = 2$ 及 f 可逆, 我们有 $f^{-1}(2) = \pi$, 故 $g(2) = \sin\left(f^{-1}(2)\right) = \sin(\pi) = 0$. 此外, 根据链式求导法则及之前关于 $\left(f^{-1}\right)'(x)$ 的方框公式, 有

$$g'(x) = \cos(f^{-1}(x)) \times (f^{-1})'(x) = \cos(f^{-1}(x)) \times \frac{1}{f'(f^{-1}(x))}.$$

将 $x = 2$ 代入且由 $f^{-1}(2) = \pi$ 及 $f'(\pi) = 5$, 得到

$$g'(2) = \cos(f^{-1}(2)) \times \frac{1}{f'(f^{-1}(2))} = \cos(\pi) \times \frac{1}{f'(\pi)} = -1 \times \frac{1}{5} = -\frac{1}{5}.$$

所以请确保你了解了之前两种形式的反函数的导数公式!

10.1.4　一个综合性例子

最后让我们以一个例子结束本节, 它将综合用到我们目前为止在本章看到的大多数理论. 假设

$$f(x) = x^2(x-5)^3, \text{ 并且其定义域为 } [2, \infty).$$

以下是我们想要做的：

(1) 证明 f 可逆;

(2) 求出反函数 f^{-1} 的定义域和值域;

(3) 检验 $f(4) = -16$;

(4) 计算 $\left(f^{-1}\right)'(-16)$.

对于问 (1), 使用乘积法则和链式求导法则可以得到

$$f'(x) = 2x(x-5)^3 + 3x^2(x-5)^2.$$

注意到 x 和 $(x-5)^2$ 是右边两项的公因子, 因此可以将它重新写作

$$f'(x) = x(x-5)^2(2(x-5) + 3x) = x(x-5)^2(5x-10) = 5x(x-5)^2(x-2).$$

当 $x > 2$ 时 (回想一下, f 的定义域是 $[2, \infty)$), 所有这三个因子 $5x$、$(x-5)^2$ 及 $(x-2)$ 都是非负的, 因此它们的乘积也是非负的. 这样我们证明了在 $(2, \infty)$ 上 $f'(x) \geqslant 0$. 此外, 在此定义域内, 唯一一处使得 $f'(x) = 0$ 的点是 $x = 5$. 由于 f 在 $[2, \infty)$ 上连续, 10.1.1 节中的方法便证明了 f 有反函数.

让我们接着来看问 (2). 反函数 f^{-1} 的值域就是 f 的定义域, 它当然就是 $[2, \infty)$. f^{-1} 的定义域则更难求一些. 确实, f^{-1} 的定义域就是 f 的值域, 因此我们需要做些工作求出这个值域. 但这不是什么大不了的. 我们知道 f 总是递增的, 这意味着 $f(2)$ 是最低点. 也就是说, 该函数始于高度 $f(2)$, 也就是 $2^2 \times (-3)^3 = -108$, 且递增向上. 那它能上升到多高呢？当 x 变得越来越大, f 也变得越来越大 —— 它的上升是没有极限的. 这意味着, f 取到自 -108 以上的所有数, 故 f^{-1} 的定义域和 f 的值域相同, 也就是 $[-108, \infty)$.

我们还需要求解问题的后两部分. 对于问 (3), 很容易计算得出 $f(4) = -16$, 这意味着 $f^{-1}(-16) = 4$. 再来看问 (4), 如果 $y = f^{-1}(x)$, 那么我们知道

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{f'(y)} = \frac{1}{5y(y-5)^2(y-2)}.$$

当 $x = -16$ 时, 从问 (3) 可知 $y = 4$. 将它代入, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{5(4)(4-5)^2(4-2)} = \frac{1}{40}.$$

这样就求解了该问题的所有部分, 但画出 $y = x^2(x-5)^3$ 的图像会有助于我们了解刚刚究竟在这里做到了什么. 我们将在 12.3.3 节回到这个例子, 深入讨论如何画出其图像, 现在我们已经能够对其图像有个大致概念了. 让我们首先在定义域 $\mathbb{R}$ 上进行操作, 最后将其限制到 $[2, \infty)$. 以下是我们所知的.

- 为了求 y 轴截距, 将 $x = 0$ 代入; 我们得到 $y = 0^2 \times (0-5)^3 = 0$. 故 y 轴截距为 0.
- 为了求 x 轴截距, 设 $x^2(x-5)^3 = 0$; 我们求出 $x = 0$ 或 $x = 5$. 这些是 x 轴截距.
- 当 x 接近于 0 时, 量 $(x-5)^3$ 非常接近 $(-5)^3 = -125$, 故 $x^2(x-5)^3$ 应该十分接近于 $-125x^2$. 该图像应该体现这一点.
- 当 x 接近于 5 时, 我们看到 x^2 非常接近 25, 故该曲线应该表现得如同 $25 \times (x-5)^3$. 而 $y = 25 \times (x-5)^3$ 的图像就如 $y = x^3$ 的图像, 只是向右平移了 5 个单位并垂直拉伸了 25 倍. 因此, 我们应该让图像反映这些信息.

汇总起来, 我们会得到类似图 10-3 的图像 (我已经将该图像的 $x < 2$ 的部分画成了虚线; 同时还要注意到两个坐标轴的比例不同).

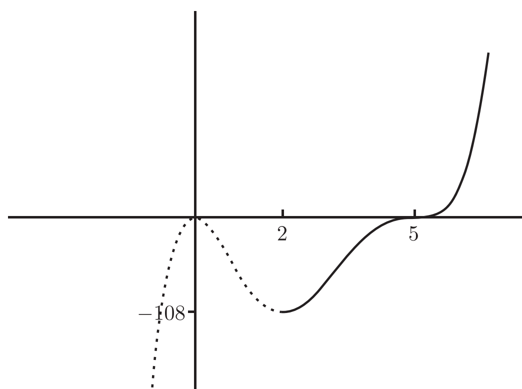


图 10-3

该图像与函数 f 在受限定义域 $[2,\infty)$ 上可逆以及 f 在此受限定义域 $[2,\infty)$ 上的值域是 $[-108,\infty)$ 等事实是一致的.

10.2 反三角函数

现在是时候来研究反三角函数了. 我们将看到如何定义反三角函数、它们的图像看上去如何以及如何对它们求导. 让我们一个个来, 首先从反正弦函数开始.

10.2.1 反正弦函数

我们先从回顾 $y=\sin(x)$ 的图像开始.

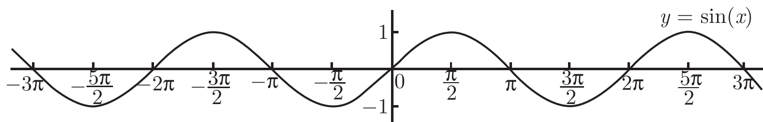


图 10-4

正弦函数有反函数吗? 从图 10-4 中可以看到, 它不满足水平线检验. 事实上, 每一条高度在 -1 和 1 之间的水平线都与图像相交**无穷**多次, 这可比我们可以容忍的零次或一次要多很多. 不过, 我们可以使用 1.2.3 节描述的方法, 尽可能少地限制定义域并使剩余部分得以通过水平线检验. 有很多选择, 但一个明智的选择是将定义域限制为区间 $[-\pi/2,\pi/2]$. 其效果如图 10-5 所示.

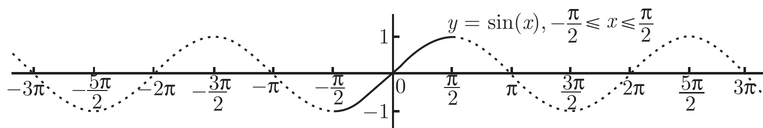


图 10-5

该曲线的实线部分就是限制定义域后所剩下的. 很明显, 我们不能往右超出 $\pi/2$, 否则随着曲线掉头往下我们又会开始重复曲线上 $\pi/2$ 左侧的值. 在 $-\pi/2$ 处也有类似的情况. 因此, 我们被困在了这个区间里.

这样如果 $f(x) = \sin(x)$, 并且其定义域为 $[-\pi/2, \pi/2]$, 则它满足水平线检验, 故它有反函数 f^{-1}. 我们将 $f^{-1}(x)$ 写成 $\sin^{-1}(x)$ 或 $\arcsin(x)$. (注意：第一个记号初看上去会有点让人困惑, 但 $\sin^{-1}(x)$ 和 $(\sin(x))^{-1}$ 不是一回事, 尽管我们有 $\sin^2(x) = (\sin(x))^2$ 及 $\sin^3(x) = (\sin(x))^3$.)

那么反正弦函数的定义域是什么呢？由于 $f(x) = \sin(x)$ 的值域是 $[-1, 1]$, 其反函数的定义域就是 $[-1, 1]$. 又由于函数 f 的定义域是 $[-\pi/2, \pi/2]$ (因为我们把定义域限制成了这样), 其反函数的值域就是 $[-\pi/2, \pi/2]$.

$y = \sin^{-1}(x)$ 的图像又如何呢？我们只需要取受限的 $y = \sin(x)$ 的图像并将它关于镜子 $y = x$ 作反射, 如图 10-6 所示. 这里有一个简洁的方法来记住该如何画这个图像. 首先, 将 $y = \sin(x)$ 的全部图像关于直线 $y = x$ 作反射, 然后抛弃图像中的其他部分, 只剩下正确部分. 图 10-7 显示了以上 $y = \sin^{-1}(x)$ 的图像是如何从翻转后的 $y = \sin(x)$ 图像中部分截取的. 注意到, 由于 $\sin(x)$ 是 x 的奇函数, 故 $\sin^{-1}(x)$ 也是如此. 这与以上图像是一致的.

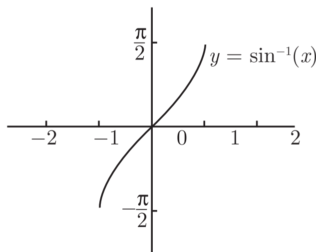


图 10-6

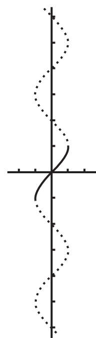


图 10-7

现在, 来对反正弦函数求导. 设 $y = \sin^{-1}(x)$, 我们想要求 $\mathrm{d}y/\mathrm{d}x$. 一种方法是写出 $x = \sin(y)$, 然后对两边关于 x 进行隐函数求导：

$$\frac{\mathrm{d}}{\mathrm{d}x}(x) = \frac{\mathrm{d}}{\mathrm{d}x}(\sin(y)).$$

左边就是 1, 而右边需要使用链式求导法则. 你应该检验一下是否得到了 $\cos(y)$ $(\mathrm{d}y/\mathrm{d}x)$. 因此, 有

$$1 = \cos(y)\frac{\mathrm{d}y}{\mathrm{d}x},$$

而它可化简为

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\cos(y)}.$$

事实上, 可以使用 10.1.3 节中的公式马上直接写出上式. 现在, 我们想要用 x 而不是 y 表示的导数. 这不成问题 —— 我们知道 $\sin(y) = x$, 因此求 $\cos(y)$ 应该不会太难. 事实上, $\cos^2(y) + \sin^2(y) = 1$, 因此 $\cos(y) = \pm\sqrt{1-x^2}$, 进而

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \pm\frac{1}{\sqrt{1-x^2}}.$$

但要选哪一个呢? 是正的还是负的? 如果你仔细观察 $y = \sin^{-1}(x)$ 的图像, 就会发现其斜率总为正. 这意味着, 我们必须取正的平方根:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\sin^{-1}(x) = \frac{1}{\sqrt{1-x^2}}, \quad \text{其中 } -1 < x < 1.}$$

注意到 $\sin^{-1}(x)$ 在端点 $x = 1$ 和 $x = -1$ 处不可导 (甚至是在单侧导数的意义下), 因为分母 $\sqrt{1-x^2}$ 在这两种情况下均为 0.

除了其导数公式和图像, 以下是关于反正弦函数其他重要事实的总结:

$$\boxed{\sin^{-1} \text{ 是奇函数; 其定义域为 } [-1,1], \text{ 值域为 } [-\pi/2, \pi/2].}$$



有了这个新的导数公式, 你应该很容易将它与乘积法则、 商法则及链式求导法则结合使用. 例如,

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sin^{-1}(7x)) \quad \text{和} \quad \frac{\mathrm{d}}{\mathrm{d}x}(x\sin^{-1}(x^3)) \text{ 是多少?}$$

对于第一个问题, 可以使用链式求导法则, 设 $t = 7x$; 或者使用 7.2.1 节结尾部分的原理: 当用 ax 替换 x 时, 你必须用 a 和导数相乘. 因此, 有

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sin^{-1}(7x)) = 7 \times \frac{1}{\sqrt{1-(7x)^2}} = \frac{7}{\sqrt{1-49x^2}}.$$

对于第二个问题, 首先我们设 $y = x\sin^{-1}(x^3)$; 此外, 将 $u = x$ 及 $v = \sin^{-1}(x^3)$ 代入, 结果是 $y = uv$. 我们需要使用乘积法则, 得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x} = \sin^{-1}(x^3) \times 1 + x\frac{\mathrm{d}v}{\mathrm{d}x}.$$

为了完成求解, 我们必须求出 $\mathrm{d}v/\mathrm{d}x$. 由于 $v = \sin^{-1}(x^3)$, 如果设 $t = x^3$, 那么 $v = \sin^{-1}(t)$. 根据链式求导法则,

$$\frac{\mathrm{d}v}{\mathrm{d}x} = \frac{\mathrm{d}v}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{1}{\sqrt{1-t^2}}(3x^2) = \frac{3x^2}{\sqrt{1-(x^3)^2}} = \frac{3x^2}{\sqrt{1-x^6}}.$$

将它代入上一个方程得到

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \sin^{-1}(x^3) \times 1 + x\frac{\mathrm{d}v}{\mathrm{d}x} = \sin^{-1}(x^3) + \frac{3x^3}{\sqrt{1-x^6}}.$$

这样我们就完成了求解.

10.2.2　反余弦函数

为了理解反余弦函数, 我们需要重复上一节的过程. 先从 $y = \cos(x)$的图像开始, 如图 10-8 所示.

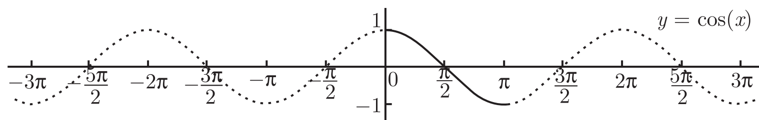


图　10-8

再一次地, 我们看到它不存在反函数. 但这次, 将定义域限制为 $[-\pi/2, \pi/2]$ 也不行, 因为在那里还是不满足水平线检验并且我们还舍弃了一部分本来有用的区间. 在上图中, 你可以看出介于 $[0, \pi]$ 的部分已经被标为实线, 这部分满足水平线检验, 因此正是我们要使用的. 这样我们得到了一个反函数, 并将它写为 $\cos^{-1}$ 或 arccos. 像反正弦函数一样, 反余弦函数的定义域是 $[-1, 1]$, 因为那是余弦函数的值域. 另一方面, 反余弦函数的值域是 $[0, \pi]$, 因为那是我们使用的余弦函数的受限定义域. $y = \cos^{-1}(x)$的图像 (图 10-9) 是通过 $y = \cos(x)$ 关于镜子 $y = x$ 反射形成的. 注意到该图像表明 $\cos^{-1}$ 既不是偶函数也不是奇函数, 尽管 $\cos(x)$ 是 x 的偶函数! 不管怎样, 如果你记不太起该如何去画以上图像, 那么可以先画出翻转后的 $\cos(x)$ 的图像, 然后选取 $[0, \pi]$ 上的那部分, 如图 10-10 所示.

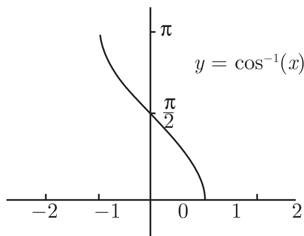


图　10-9

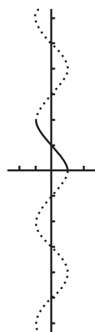


图　10-10

现在是时候来对 $y = \cos^{-1}(x)$ 关于 x 求导了. 我们要做与上一节完全相同的操作. 首先写出 $x = \cos(y)$ 并对它关于 x 进行隐函数求导：

$$\frac{\mathrm{d}}{\mathrm{d}x}(x) = \frac{\mathrm{d}}{\mathrm{d}x}(\cos(y)).$$

左边是 1, 右边是 $-\sin(y)\,(\mathrm{d}y/\mathrm{d}x)$. 重新整理可得

$$\frac{dy}{dx} = -\frac{1}{\sin(y)}.$$

由于 $\cos^2(y) + \sin^2(y) = 1$, 且 $x = \cos(y)$, 我们有 $\sin(y) = \pm\sqrt{1-x^2}$. 这意味着,

$$\frac{dy}{dx} = -\frac{1}{\pm\sqrt{1-x^2}} = \pm\frac{1}{\sqrt{1-x^2}}.$$

不同于反正弦函数, 反余弦函数的图像是向下的, 这意味着, 其斜率总为负, 因此得到

$$\boxed{\frac{d}{dx}\cos^{-1}(x) = -\frac{1}{\sqrt{1-x^2}}, \quad \text{其中 } -1 < x < 1.}$$

下面则是我们前面发现的关于反余弦函数的其他一些事实：

$\cos^{-1}$ 既不是偶函数也不是奇函数; 其定义域为 $[-1,1]$, 值域为 $[0,\pi]$.

在转入讨论反正切函数之前, 让我们试着将反正弦函数和反余弦函数的导数并排放在一起：

$$\frac{d}{dx}\sin^{-1}(x) = \frac{1}{\sqrt{1-x^2}} \quad \text{和} \quad \frac{d}{dx}\cos^{-1}(x) = -\frac{1}{\sqrt{1-x^2}}.$$

这两个导数互为相反数! 让我们试着来看一下为什么这说得通. 如果你将 $y = \sin^{-1}(x)$ 和 $y = \cos^{-1}(x)$ 画在同一坐标系中, 会得到图 10-11.

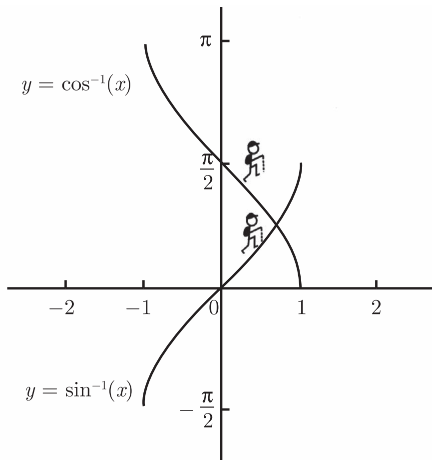


图 10-11

在同一条垂线上, 图中的两个登山者面对的情况恰恰相反, 因此这两个导数互为相反数是说得通的. 确实, 我们现在知道

$$\frac{d}{dx}(\sin^{-1}(x) + \cos^{-1}(x)) = \frac{1}{\sqrt{1-x^2}} - \frac{1}{\sqrt{1-x^2}} = 0.$$

因此 $y=\sin^{-1}(x)+\cos^{-1}(x)$ 的斜率为常数 0, 这意味着它始终是平的. 事实上, 如果将上图中这两个函数值的高度相加, 你会看到对于任意的值 x 都会得到 $\pi/2$. 这样我们刚刚使用微积分证明了以下恒等式：对于在区间 $[-1,1]$ 上任意的 x,

$$\sin^{-1}(x)+\cos^{-1}(x)=\frac{\pi}{2}.$$

稍作思考, 你就会意识到这是说得通的! 看一下图 10-12.

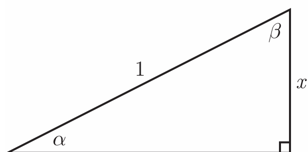


图　10-12

由于 $\sin(\alpha)=x$, 我们有 $\alpha=\sin^{-1}(x)$. 类似地, $\cos(\beta)=x$ 意味着 $\beta=\cos^{-1}(x)$. 又由于 $\alpha+\beta=\pi/2$, 这就意味着

$$\sin^{-1}(x)+\cos^{-1}(x)=\frac{\pi}{2}.$$

看到微积分与几何学殊途同归, 这很棒, 不是吗?

10.2.3　反正切函数

下面我们继续. 让我们先回忆一下 $y=\tan(x)$ 的图像, 如图 10-13 所示.

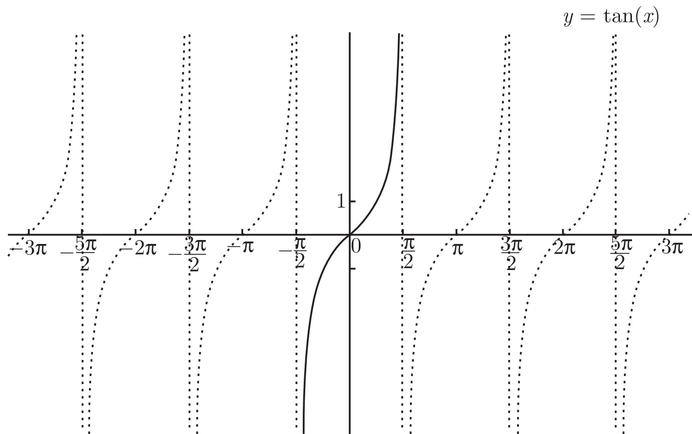


图　10-13

我们将定义域限制在 $(-\pi/2,\pi/2)$, 以便可以得到反函数 $\tan^{-1}$, 也可写作 arctan. 反函数的定义域是正切函数的值域, 即所有的 $\mathbb{R}$. 它的值域是 $(-\pi/2,\pi/2)$, 自然也就是我们现在所取的 $\tan(x)$ 的受限定义域. $y=\tan^{-1}(x)$ 的图像如图 10-14 所示.

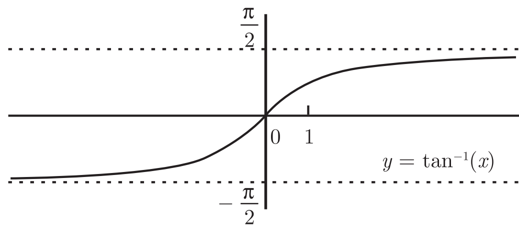


图　10-14

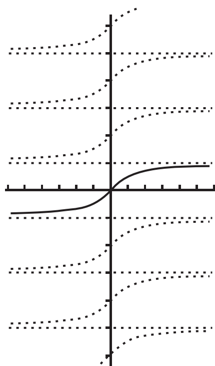


图 10-15



从图像上可见, $\tan^{-1}(x)$ 是 x 的奇函数 —— 事实上, 它继承了 $\tan(x)$ 的奇函数性质. 再一次地, 通过翻转 $y = \tan(x)$ 的图像并将大部分删除, 你就可以记起它的图像, 如图10-15所示.

现在, 让我们来对 $y = \tan^{-1}(x)$ 关于 x 求导. 写出 $x = \tan(y)$ 并对它关于 x 进行隐函数求导. 自己验证一下,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\sec^2(y)}.$$

由于 $\sec^2(y) = 1+\tan^2(y)$, 且 $\tan(y) = x$, 我们有 $\sec^2(y) = 1 + x^2$. 这意味着

$$\boxed{\text{对于所有的实数 } x,\ \frac{\mathrm{d}}{\mathrm{d}x}\tan^{-1}(x) = \frac{1}{1+x^2}.}$$

从之前的分析中我们还得知：

$$\boxed{\tan^{-1} \text{ 是奇函数; 其定义域是 } \mathbb{R} \text{ 且值域是 } (-\pi/2, \pi/2).}$$

不同于反正弦函数和反余弦函数, 反正切函数有水平渐近线. (前两个函数根本没机会有, 因为它们的定义域都是 $[-1, 1]$.) 从其图像可以看到, 当 $x \to \infty$ 时 $\tan^{-1}(x)$ 趋于 $\pi/2$, 而当 $x \to -\infty$ 时 $\tan^{-1}(x)$ 趋于 $-\pi/2$. 事实上, 正切函数在 $x = \pi/2$ 和 $x = -\pi/2$ 处的垂直渐近线变成了反正切函数的水平渐近线. 这意味着, 我们有以下有用的极限：

$$\boxed{\lim_{x\to\infty} \tan^{-1}(x) = \frac{\pi}{2}} \quad \text{和} \quad \boxed{\lim_{x\to-\infty} \tan^{-1}(x) = -\frac{\pi}{2}.}$$



顺便说一下, 我们在 3.5 节其实已经看到过这些极限. 不管怎样, 当其他形式的虚拟变量也趋于 $\pm\infty$ 时, 这两个极限依然成立. 例如, 为了求

$$\lim_{x\to-\infty} \frac{x^2 - 6x + 4}{(2x^2 + 7x - 8)\tan^{-1}(3x)},$$

我们先将分式拆开, 得到

$$\lim_{x\to-\infty} \frac{x^2 - 6x + 4}{2x^2 + 7x - 8} \times \frac{1}{\tan^{-1}(3x)}.$$

第一个分式的极限是 $1/2$, (自己验证一下!) 第二个分式呢? 好吧, 当 x 在负方向上变得非常大时, $3x$ 也一样, 故 $\tan^{-1}(3x)$ 趋于 $-\pi/2$. 因此, 整个极限是

$$\frac{1}{2} \times \frac{1}{-\frac{\pi}{2}} = -\frac{1}{\pi}.$$



然而, 假若我们用 $3x^2$ 替换 $3x$：

$$\lim_{x\to-\infty}\frac{x^2-6x+4}{(2x^2+7x-8)\tan^{-1}(3x^2)}.$$

现在当 $x\to-\infty$ 时, $\tan^{-1}\left(3x^2\right)$ 趋于 $\pi/2$, 因为 $3x^2$ 趋于 ∞, 而不是 $-\infty$. 因此在这种情况下, 整个极限是 $1/\pi$.

10.2.4　反正割函数

旅程继续. 图 10-16 是 $y=\sec(x)$ 的图像.

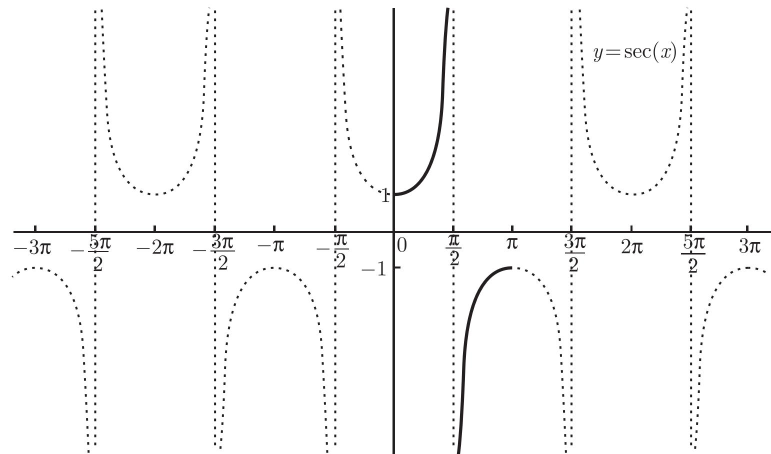


图　10-16

现在的情况与我们求反余弦函数时的情况 (不出意外地) 非常相似. 我们必须将定义域限制在 $[0,\pi]$ 上, 并除去点 $\pi/2$, 因为它甚至不在 $\sec(x)$ 的原始定义域中. 正割函数的值域是 $(-\infty,-1]$ 和 $[1,\infty)$ 这两个区间的并集, 因此这也是其反函数 $\sec^{-1}$(或 arcsec) 的定义域. 至于 $\sec^{-1}$ 的值域, 它和原函数的受限定义域是一样的：$[0,\pi]$ 除去点 $\pi/2$. 它的图像如图 10-17 所示.

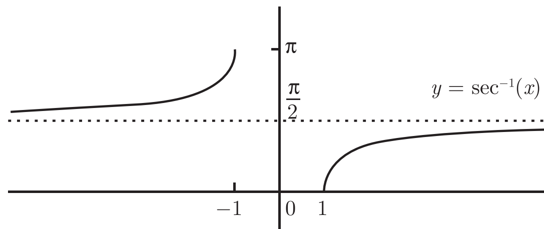


图　10-17

注意到在 $y=\pi/2$ 处, 有一条双侧水平渐近线, 故

$$\boxed{\lim_{x\to\infty} \sec^{-1}(x) = \frac{\pi}{2}} \quad 和 \quad \boxed{\lim_{x\to-\infty} \sec^{-1}(x) = \frac{\pi}{2}.}$$

现在让我们来求导吧. 如果 $y = \sec^{-1}(x)$, 那么 $x = \sec(y)$, 故

$$\frac{\mathrm{d}}{\mathrm{d}x}(x) = \frac{\mathrm{d}}{\mathrm{d}x}(\sec(y)).$$

要确保你知道为什么会有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\sec(y)\tan(y)}.$$

由于 $x = \sec(y)$, 又由于 $\sec^2(y) = 1 + \tan^2(y)$, 我们可以重新整理并取平方根, 得到 $\tan(y) = \pm\sqrt{x^2-1}$. 这意味着

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\pm x\sqrt{x^2-1}}.$$

那它是正的还是负的呢? 回头看一下 $y = \sec^{-1}(x)$ 的图像, 你可以看到其斜率总为正. 因此, 我们实际上需要做得更聪明些 —— 不是简单取正号或负号, 而是我们用 $|x|$ 代替 x, 这样我们总能得到正的结果. 也就是说,

$$\boxed{对于\ x > 1 \quad 或 \quad x < -1,\ \frac{\mathrm{d}}{\mathrm{d}x}\sec^{-1}(x) = \frac{1}{|x|\sqrt{x^2-1}}.}$$

有关反正割函数的其他事实可以总结如下:

> $\sec^{-1}$ 既不是奇函数也不是偶函数; 其定义域是 $(-\infty,-1]\cup[1,\infty)$ 且值域是 $[0,\pi]\setminus\left\{\frac{\pi}{2}\right\}$.

(这里我用标准缩写 $\cup$ 来表示两个区间的并集, 用 $\setminus$ 表示 “不包括”.)

10.2.5 反余割函数和反余切函数

现在让我们快速看一下最后两个反三角函数. 你可以重复之前的分析来求得 $y = \csc^{-1}(x)$ 和 $y = \cot^{-1}(x)$ 的定义域、值域以及图像:

> $\csc^{-1}$ 是奇函数; 其定义域为 $(-\infty,-1]\cup[1,\infty)$ 且值域是 $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]\setminus\{0\}$.

> $\cot^{-1}$ 既不是奇函数也不是偶函数; 其定义域为 $\mathbb{R}$ 且值域是 $(0,\pi)$.

图 10-18 是它们的图像.

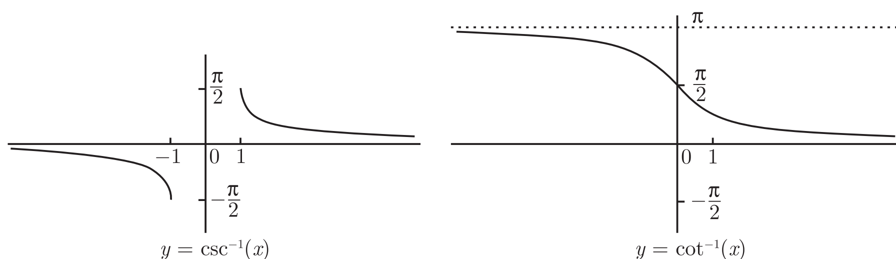


图　10-18

这两个函数都有水平渐近线：$y = \csc^{-1}(x)$ 在 $y = 0$ 处有一条双侧水平渐近线, $y = \cot^{-1}(x)$ 在 $y = \pi$ 处有一条左侧水平渐近线, 而在 $y = 0$ 处有一条右侧水平渐近线. 我们可以将这些极限总结如下：

$$\boxed{\lim_{x\to\infty} \csc^{-1}(x) = 0} \quad \text{和} \quad \boxed{\lim_{x\to-\infty} \csc^{-1}(x) = 0.}$$

$$\boxed{\lim_{x\to\infty} \cot^{-1}(x) = 0} \quad \text{和} \quad \boxed{\lim_{x\to-\infty} \cot^{-1}(x) = \pi.}$$

当然, 如果你知道上述图像, 就可以很容易地写出这些极限, 而不需要死记硬背. 注意到 $y = \csc^{-1}(x)$ 的图像和 $y = \sec^{-1}(x)$ 的图像非常相似; 事实上, 你可以通过将一个图像关于 $y = \pi/4$ 作对称得到另一个图像. 这正好与 $y = \sin^{-1}(x)$ 和 $y = \cos^{-1}(x)$ 的相互关系是一致的. 因此毫不奇怪, $\csc^{-1}(x)$ 的导数就是负的 $\sec^{-1}(x)$ 的导数：

$$\boxed{\text{对于 } x > 1 \text{ 或 } x < -1,\ \frac{\mathrm{d}}{\mathrm{d}x}\csc^{-1}(x) = -\frac{1}{|x|\sqrt{x^2-1}}.}$$

同样的事情也见于 $\cot^{-1}(x)$ 和 $\tan^{-1}(x)$, 因此

$$\boxed{\text{对于所有的实数 } x,\ \frac{\mathrm{d}}{\mathrm{d}x}\cot^{-1}(x) = -\frac{1}{1+x^2}.}$$

10.2.6　计算反三角函数

这样我们就完成了对于反三角函数一个相当全面的讨论. 由于我们又多了几个求导法则, 练习一下对涉及反三角函数的函数求导似乎是个不错的主意. 不过同时, 我们也不该忽略反三角函数一些不涉及任何微积分但基本的计算. 首先说, 你应该确保自己不费力就可以算出诸如 $\sin^{-1}(1/2)$、$\cos^{-1}(1)$ 以及 $\tan^{-1}(1)$ 等的值. 例如, 为了求 $\sin^{-1}(1/2)$, 你要立马能想到自己是要在 $[-\pi/2, \pi/2]$ 上找一个角, 其正弦值是 $1/2$. 而这当然就是 $\pi/6$. 类似地, 你也应该抬笔就可以写出 $\cos^{-1}(1) = 0$ 和 $\tan^{-1}(1) = \pi/4$. 所有这些常用值都列在了 2.1 节结尾部分的那张表里.



下面是一个更有趣的问题：该如何化简

$$\sin^{-1}\left(\sin\left(\frac{13\pi}{10}\right)\right)?$$

本能的反应是消去反正弦函数和正弦函数, 只剩下 $13\pi/10$. 但这显然不对, 因为正如我们在 10.2.1 节看到的, 反正弦函数的值域是 $[-\pi/2,\pi/2]$. 因此, 我们需要做的就是找到在那个区间里的一个角, 其正弦值与 $13\pi/10$ 的正弦值一样. 好吧, 注意到 $13\pi/10$ 在第三象限 (因为它大于 π 但小于 $3\pi/2$), 因此它的正弦值是负的. 此外, 其参照角是 $3\pi/10$. 在 $[-\pi/2,\pi/2]$ 中带有相同参照角的角是 $3\pi/10$ 和 $-3\pi/10$. 前者的正弦值为正, 后者的正弦值为负. 而我们需要一个负的正弦值, 因此我们证明了

$$\sin^{-1}\left(\sin\left(\frac{13\pi}{10}\right)\right)=-\frac{3\pi}{10}.$$

那么又该如何化简

$$\cos^{-1}\left(\cos\left(\frac{13\pi}{10}\right)\right)?$$

之前的答案 $-3\pi/10$ 在这里又显然不对, 因为反余弦函数的值域是 $[0,\pi]$. 哎, 为什么就这么麻烦呢? 很遗憾, 我也无能为力, 事情就是这样 …… 因此, 让我们再来做一遍: $13\pi/10$ 在第三象限, 因此其余弦值为负; 其参照角是 $3\pi/10$, 而在 $[0,\pi]$ 中带有相同参照角的角是 $3\pi/10$ 和 $7\pi/10$; 这两个角的余弦值分别是正的和负的, 而我们想要一个负的余弦, 因此我们必须有

$$\cos^{-1}\left(\cos\left(\frac{13\pi}{10}\right)\right)=\frac{7\pi}{10}.$$

现在, 我留给你来证明

$$\tan^{-1}\left(\tan\left(\frac{13\pi}{10}\right)\right)=\frac{3\pi}{10}.$$

只需回想一下, 正切函数在第三象限为正! 不管怎样, 这些都是很难的例子, 因此如果你认为求

$$\sin\left(\sin^{-1}\left(-\frac{1}{5}\right)\right)$$

也很难的话, 我不会责怪你. 幸运的是, 它不难, 答案就是 $-\frac{1}{5}$. 一般地, $\sin\big(\sin^{-1}(x)\big) = x$, 只要 x 在反正弦函数的定义域 $[-1,1]$ 中. (否则的话, $\sin\left(\sin^{-1}(x)\right)$ 甚至都说不通!) 但当你试图写出 $\sin^{-1}(\sin(x))=x$ 时, 问题就出现了, 因为这根本不对, 前面 $13\pi/10$ 的例子就说明了这一点. 当然, 同样的情况对其他反三角函数也成立. (参见 1.2.4 节结尾部分的讨论.)

再来看两个例子: 考虑该如何求

$$\sin\left(\cos^{-1}\left(\frac{\sqrt{15}}{4}\right)\right)\quad 和 \quad \sin\left(\cos^{-1}\left(-\frac{\sqrt{15}}{4}\right)\right).$$

求解这两种情况的技巧是, 使用三角恒等式 $\cos^2(x)+\sin^2(x)=1$. 对于第一个问题, 令

$$x=\cos^{-1}\left(\frac{\sqrt{15}}{4}\right),$$

并注意到我们想要求 $\sin(x)$. 我们实际上知道 $\cos(x)$:

$$\cos(x)=\cos\left(\cos^{-1}\left(\frac{\sqrt{15}}{4}\right)\right)=\frac{\sqrt{15}}{4}.$$

回想一下, 取一个反余弦的余弦构不成问题：反过来才有可能出现问题. 不管怎样, 我们知道 $\cos(x)$, 因此通过重新整理恒等式 $\cos^2(x)+\sin^2(x)=1$, 我们必须有

$$\sin(x)=\pm\sqrt{1-\cos^2(x)}=\pm\sqrt{1-\left(\frac{\sqrt{15}}{4}\right)^2}=\pm\sqrt{\frac{1}{16}}=\pm\frac{1}{4}.$$

因此, 我们想要的答案是 $1/4$ 或 $-1/4$. 但到底是哪一个呢? 由于 $\sqrt{15}/4$ 是正的, 它的反余弦必定位于 $[0,\pi/2]$. 也就是说, x 在第一象限, 故其正弦为正. 最终, 我们证明了

$$\sin\left(\cos^{-1}\left(\frac{\sqrt{15}}{4}\right)\right)=\frac{1}{4}.$$

至于

$$\sin\left(\cos^{-1}\left(-\frac{\sqrt{15}}{4}\right)\right),$$

你可以重复上述过程来证明

$$\sin(x)=\pm\sqrt{1-\cos^2(x)}=\pm\sqrt{1-\left(-\frac{\sqrt{15}}{4}\right)^2}=\pm\sqrt{\frac{1}{16}}=\pm\frac{1}{4}.$$

你可能会猜这次的答案该是 $-1/4$, 但这是乱猜. 你看, $-\sqrt{15}/4$ 是负的, 故其反余弦必定位于区间 $[\pi/2,\pi]$. 也就是说, x 在第二象限. 但正弦函数在第二象限还是正的! 因此, $\sin(x)$ 必定为正, 这样我们也就证明了

$$\sin\left(\cos^{-1}\left(-\frac{\sqrt{15}}{4}\right)\right)=\frac{1}{4}.$$

事实上, 我们可以注意到, $\sin\left(\cos^{-1}(A)\right)$ 必定总是非负的, 尽管如果 A 是负的 (注意到 A 必须位于 $[-1,1]$, 因为那是反余弦函数的定义域). 这是因为 $\cos^{-1}(A)$ 位于区间 $[0,\pi]$, 正弦函数在这个区间上是非负的.

事实上, 在后面 19.3 节讨论三角换元法时, 我们会再来看另一种求解形如 $\sin\left(\cos^{-1}(A)\right)$ 的方法. 不过现在, 暂且告别这些反三角函数, 并快速看一下反双曲函数.

10.3 反双曲函数

双曲函数 (参见 9.7 节) 的情况有点不同. 回想一下这些函数的图像是什么样的. 特别地, 你可以看到 $y = \cosh(x)$ 的图像有点像 $y = x^2$ 的图像, 只是向上移动了 1 且形状略有不同. 如果想求这个函数的反函数, 必须舍弃该图像的左半部分, 就像为了取正的平方根而舍弃负的那个一样. 另一方面, $y = \sinh(x)$ 已经满足水平线检验, 因此没有必要再做什么. 这样, 我们得到带有以下性质的两个反函数:

$\cosh^{-1}$ 既不是奇函数也不是偶函数; 其定义域是 $[1, \infty)$ 且值域是 $[0, \infty)$.

$\sinh^{-1}$ 是奇函数; 其定义域和值域都是 $\mathbb{R}$.

像往常一样, 它们的图像可以通过将原始图像关于直线 $y = x$ 反射获得, 如图 10-19 所示.

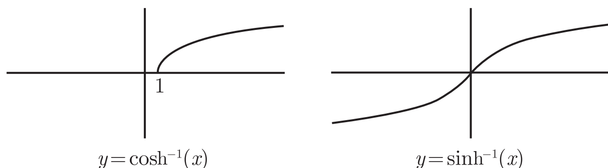


图 10-19

我们可以使用与求反三角函数的导数相同的方法来求它们的导数. 特别地, 如果 $y = \cosh^{-1}(x)$, 那么 $x = \cosh(y)$; 对它关于 x 作隐函数求导, 我们得到

$$1 = \sinh(y)\frac{\mathrm{d}y}{\mathrm{d}x}.$$

(回想一下, $\cosh(x)$ 关于 x 的导数是 $\sinh(x)$, 而不是 $-\sinh(x)$.) 又由于 $\cosh^2(x) - \sinh^2(x) = 1$, 因此我们可以对它重新整理并取平方根, 得到 $\sinh(y) = \pm\sqrt{\cosh^2(y) - 1} = \pm\sqrt{x^2 - 1}$. 由于 $\cosh^{-1}(x)$ 在 x 上明显是递增的, 故我们有

$$\text{对于 } x > 1,\ \frac{\mathrm{d}}{\mathrm{d}x}\cosh^{-1}(x) = \frac{1}{\sqrt{x^2 - 1}}.$$

以相同的方法, 你应该可以得到

$$\text{对于所有的实数 } x,\ \frac{\mathrm{d}}{\mathrm{d}x}\sinh^{-1}(x) = \frac{1}{\sqrt{x^2 + 1}}.$$

现在, 暂且把微积分放在一边, 让我们回想一下 $\cosh(x)$ 和 $\sinh(x)$ 的定义:

$$\cosh(x) = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2} \quad \text{和} \quad \sinh(x) = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2}.$$

由于我们可以用指数函数来表示 $\cosh(x)$ 和 $\sinh(x)$, 因而应该也可以用对数函数来表示反函数, 毕竟指数函数和对数函数互为反函数. 让我们来看一下这是如何做

到的. 例如, 如果 $y=\cosh^{-1}(x)$, 那么 $x=\cosh(y)=(\mathrm{e}^y+\mathrm{e}^{-y})/2$. 现在, 你可以用一个小技巧来求解 y. 令 $u=\mathrm{e}^y$, 那么 $\mathrm{e}^{-y}=1/u$. 方程变为

$$x=\frac{u+1/u}{2}.$$

两边同乘以 $2u$ 并整理得到一个 u 的二次方程：$u^2-2xu+1=0$. 根据二次公式,

$$\mathrm{e}^y=u=x\pm\sqrt{x^2-1},$$

然后对两边取对数,

$$y=\ln(x\pm\sqrt{x^2-1}).$$

那么到底取正号还是负号呢? 略作思考后可以看到, 如果 $x>1$, 那么 $x-\sqrt{x^2-1}<1$. 这意味着, 它的对数是负的. (回想一下, 一个介于 0 和 1 的数的对数是负的!) 这不是我们想要的. 因此, 它是正的平方根. 这样我们证明了, 当 $x\geqslant 1$ 时,

$$\cosh^{-1}(x)=\ln(x+\sqrt{x^2-1}).$$

类似地, 你可以证明, 对于所有的 x,

$$\sinh^{-1}(x)=\ln(x+\sqrt{x^2+1}).$$



作为练习, 你应该尝试对这后两个方程的右边求导并检验你的答案是否与我们之前求出的 $\cosh^{-1}(x)$ 和 $\sinh^{-1}(x)$ 的导数一致.

其他的反双曲函数

到目前为止, 我们只研究了双曲正弦函数和双曲余弦函数的反函数. 如果你对其他四个双曲函数重复这个分析过程, 应该可以得出以下结论.

$\tanh^{-1}$ 是奇函数; 其定义域是 $(-1,1)$, 值域是 $\mathbb{R}$.

sech^{-1} 既不是奇函数也不是偶函数; 其定义域是 $(0,1]$, 值域是 $[0,\infty)$.

csch^{-1} 是奇函数; 其定义域和值域都是 $\mathbb{R}\backslash\{0\}$.

$\coth^{-1}$ 是奇函数; 其定义域是 $(-\infty,-1)\cup(1,\infty)$, 值域是 $\mathbb{R}\backslash\{0\}$.

注意到为了得到反函数, 我们已经将 sech 的定义域限制为 $[0,\infty)$, 正如我们对 cosh 所做的那样.

图 10-20 是它们的图像, 试与 9.7 节中原函数的图像作比较.

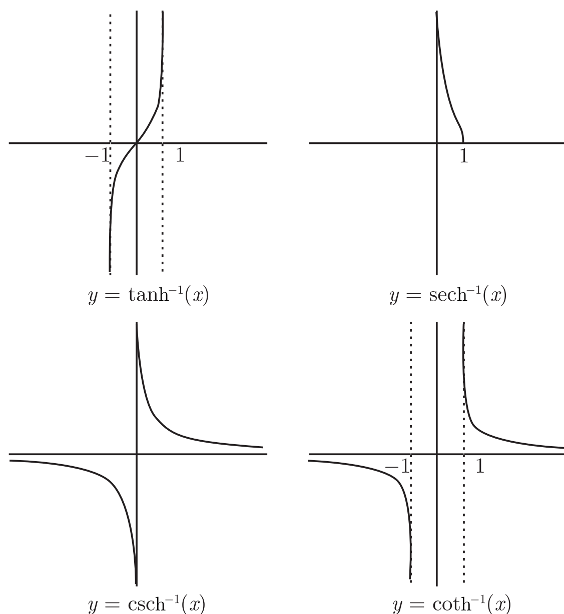


图 10-20

最后, 通过整理出 x 的式子并对它使用关于 x 进行隐函数求导的标准技巧, 你就可以求其导数. 下面就是各自的导数：

$$\frac{\mathrm{d}}{\mathrm{d}x}\tanh^{-1}(x) = \frac{1}{1-x^2} \quad (-1 < x < 1)$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\coth^{-1}(x) = \frac{1}{1-x^2} \quad (x > 1 \text{ 或 } x < -1)$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\text{sech}^{-1}(x) = -\frac{1}{x\sqrt{1-x^2}} \quad (0 < x < 1)$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\text{csch}^{-1}(x) = -\frac{1}{|x|\sqrt{1+x^2}} \quad (x \neq 0).$$

要记住, 所有这些导数公式只有当 x 在相关函数本身的定义域内时才成立. 这就解释了为什么 $\tanh^{-1}(x)$ 和 $\coth^{-1}(x)$ 的导数是相同的, 尽管它们的图像看起来非常不同. 具体说, $\tanh^{-1}(x)$ 只有在 $(-1,1)$ 上有定义, 而 $\coth^{-1}(x)$ 只有在区间 $[-1,1]$ 之外有定义. 它们没有重叠部分, 因此这两个函数有相同的导数不会造成问题. 就这样吧, 反函数已经讨论得够多了!

第 11 章　导数和图像

我们已经看过了怎样求导不同类型的函数：多项式和多项式型函数、三角函数和反三角函数、指数函数和对数函数, 以及双曲函数和反双曲函数. 现在我们可以利用这些知识来绘制一般函数的图像. 我们将看到导数会如何帮助我们理解函数的最大值和最小值, 而二阶导数又会如何帮助我们理解函数所谓的凹性. 总的来说, 我们要介绍以下知识点：

(1) 函数的局部和全局极值问题, 以及怎样用导数去找极值;

(2) 罗尔定理和中值定理, 以及它们对绘制函数图像的意义;

(3) 二阶导数的图像阐释;

(4) 对导数为零点的分类.

在下一章中, 我们将看到一种借助上述手段绘制函数图像的综合性方法.

11.1　函数的极值

如果我们说 $x = a$ 是函数 f 的一个**极值点**, 这就意味着函数 f 在 a 点处有最大值或最小值. 在 5.1.6 节中, 我们已经讨论到一点最大值和最小值; 在学习下面的内容之前, 我强烈建议你翻回第 5 章复习一下. 不管怎样, 我们需要讨论得更深入一些, 并区分两种类型的极值：全局极值和局部极值.

11.1.1　全局极值和局部极值

最大值的基本思想是, 它是函数图像的最高点. 考虑图 11-1 中函数在定义域 [0, 7] 内的最大值.

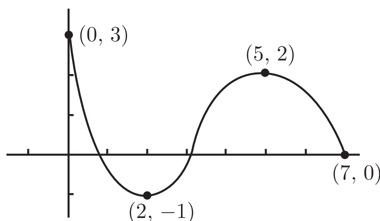


图　11-1

显然该函数达到的最高点为 3, 出现在 $x = 0$ 处, 因此你可以说该函数在 $x = 0$ 处有最大值. 另一方面, 想象这个图像为一座山的截面, 而你正在攀登它. 假设你从点 (2, −1) 开始, 往右向上攀登. 最终你到达了山峰 (5, 2), 然后你开始往下走. 这

个山峰无疑会让人感觉, 它是某种最大值 —— 它是山的顶部, 高度为 2, 尽管在它左侧还有一个山峰比它更高. 如果在 $x=0$ 处的山峰被云雾笼罩, 你在点 (5, 2) 时看不到它, 这样你会确实感觉自己是在最高点了. 事实上, 如果我们限制定义域为 [2, 7], 这时 $x=5$ 处确实有最大值.

我们需要一种方法来区分这种情况. 如果当 $x=a$ 时, $f(a)$ 是函数 f **整个**定义域内的最大值, 我们就说它是**全局最大值** (或**绝对最大值**). 用记号表示, 我们说对于在该函数定义域中的任何数值 x 都有 $f(a) \geqslant f(x)$. 这是我们之前使用过的定义, 但这次我们定义得更准确, 特称它为 "全局最大值", 而不是泛泛称 "最大值".

正如之前注意到的, 一个函数可能有多个全局最大值. 例如 $\cos(x)$ 的最大值为 1, 但有无数个 x 的值与之对应. (从 $y=\cos(x)$ 的图像中可以看到, 这些值都是 2π 的整数倍.)

那么另一类最大值又是什么情况呢? 在包含 a 的某一小段区间内, 如果在 $x=a$ 处, $f(a)$ 有最大值, 我们就称这点为**局部最大值**, 或**相对最大值**. 你可以把这想象成, 舍弃定义域的大部分, 而只关注靠近 a 的 x 的值, 然后称函数在这些 x 值中达到最大值.

让我们看看在上面的例子中这是如何运作的. 我们发现在 $x=5$ 处有局部最大值, 因为如果只关注 $x=5$ 临近部分的函数, 点 (5, 2) 就是最高点. 例如, 如果我们把图像向左延伸到 $x=3$, 点 (5, 2) 依然是最高点. 但 $x=5$ 不是全局最大值, 因为点 (0, 3) 更高. 这意味着 $x=0$ 是全局最大值. 当然, 它也是局部最大值, 事实上很明显, **每一个全局最大值都是局部最大值**.

用同样的方式, 我们也可以定义全局和局部最小值. 在上图中, 我们可以看出 $x=2$ 是全局最小值 (值为 -1), 因为它的高度最低. 另一方面, $x=7$ 是局部最小值 (值为 0). 的确, 如果你看图像右侧从 $x=5$ 到 $x=7$ 这一段, 会发现右端点 $x=7$ 就是该段的最低点.

11.1.2 极值定理

在第 5 章中, 我们看到过最大值与最小值定理. 它说的是, **连续**函数在一个**闭**区间 $[a, b]$ 内一定有一个全局最大值和一个全局最小值. 如果函数不是连续的, 或者尽管连续但其定义域不是一个闭区间, 这时该函数可能没有全局最大值或最小值. 例如, 定义在闭区间 $[-1, 1]$ 上、但 x 不能为 0 的函数 $f(x)=1/x$, 其定义域内就没有全局最大值和最小值. (画出图像找原因!)

最大值与最小值定理的问题在于, 它没有告诉我们全局最大值和最小值出现的位置. 这时导数就有了用武之地. 如果函数在 $x=c$ 处的导数为零或导数不存在, 我们就称 $x=c$ 为**临界点**. 然后我们有以下这个很好的结论[1]:

---

① 最大值与最小值定理也经常会被称为极值定理, 有时会与这里的极值定理放到一起.

极值定理　假设函数 f 定义在开区间 (a,b) 内, 并且点 c 在 (a,b) 区间内. 如果点 c 为函数的局部最大值或最小值, 那么点 c 一定为该函数的临界点. 也就是说, $f'(c)=0$ 或 $f'(c)$ 不存在.

所以在一个开区间内的局部最大值和最小值只可能出现在临界点. 但反过来说, 临界点一定是局部最大值或最小值就不一定成立. 例如, 如果函数 $f(x)=x^3$, 它的导数为 $f'(x)=3x^2$, 可以看出 $f'(0)=0$. 这意味着 $x=0$ 是该函数的临界点. 另一方面, 从图像 $y=x^3$ 中可以看出, 该点既不是局部最大值也不是局部最小值.

上述定理适用于开区间. 但如果定义域为闭区间 $[a,b]$, 情况又会怎样呢？端点 a 和 b 可能是局部最大值或最小值, 而它们不在上述定理的讨论范围内. 综上所述, 在一个闭区间内, 局部最大值或最小值只可能出现在临界点, 或该区间的端点. 例如, 让我们更仔细地看看图 11-2.



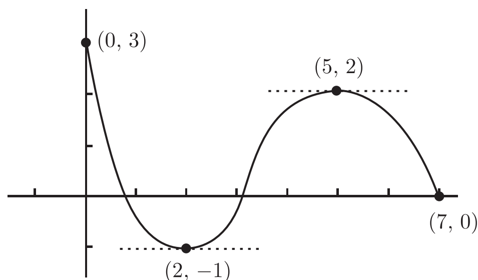


图　11-2

从图像中可以看出, 局部最大值出现在 $x=0$ 和 $x=5$ 处, 局部最小值出现在 $x=2$ 和 $x=7$ 处. 在 $x=2$ 和 $x=5$ 处的斜率为零, 所以这两点为临界点; 而 $x=0$ 和 $x=7$ 这两点为端点.

你可能会想为什么极值定理说得通. 假设在 $x=a$ 处有局部最小值, 当你从左边接近 $x=a$ 时, 你必定是在下坡, 所以斜率 (如果存在的话) 是负的. 当你从右边离开 $x=a$ 时, 你是在上坡, 所以斜率是正的. 斜率从负到正, 你自然会想到之间有斜率为零的一点. 另一方面, 如果 $f(x)=|x|$, 它的斜率从 -1 直接跳到 1, 而没有经过斜率为零的阶段. 这是因为 $f'(0)$ 不存在 (参见 5.2.10 节). 不过, 这没有关系——$x=0$ 仍是临界点, 因为在那里导数不存在. 它也是局部最小值. (你知道原因吗?) 顺便说一下, 上述推理不算严格证明, 真正的证明请参见附录 A 的 A.6.6 节.

11.1.3　求全局最大值和最小值

极值定理让求全局最大值和最小值变得轻而易举, 因为它缩小了它们可能的存身之处. 基本思路是这样的：每一个全局极值也是局部极值, 而局部极值只可能出现在临界点, 所以找出所有的临界点并求出它们对应的函数值, 这样其中最大的就

是全局最大值, 最小的就是全局最小值. 下面是怎样求在闭区间 $[a, b]$ 内的全局最大值和最小值的详细步骤.



(1) 找出 $f'(x)$, 并列出在 (a,b) 中 $f'(x)$ 不存在或 $f'(x) = 0$ 的点. 也就是说, 列出在开区间 (a, b) 内所有的临界点.

(2) 把端点 $x = a$ 和 $x = b$ 放入上述列表.

(3) 对于上述列表中的每一个点, 将它们代入 $y = f(x)$ 以求出它们所对应的函数值.

(4) 找出最大的函数值以及它所对应的 x 的值, 这就是全局最大值.

(5) 用同样的方法找到最小的函数值和全局最小值.

我们会在后面的 11.5 节再考虑局部极值, 而现在先来看一下上述方法的一个应用例子. 假设 $f(x) = 12x^5 + 15x^4 - 40x^3 + 1$, 其定义域为 $[-1, 2]$, 在此定义域内的全局最大值和最小值是什么?



让我们执行上述程序. 第一步, 需要找出 $f'(x)$. 这不成问题：你可以检验一下 $f'(x) = 60x^4 + 60x^3 - 120x^2$. 很明显, 在开区间 $(-1, 2)$ 内该函数的导数都存在, 所以我们仅仅需要去找满足导数为零的所有 x 的值. 如果对导函数进行因式分解, 得到 $f'(x) = 60x^2(x-1)(x+2)$, 我们就可以很容易找到使导函数为零的所有 x 的值：要使 $f'(x) = 0$, 必须有 $x = 0$、$x = 1$ 或 $x = -2$. 由于 -2 不在定义域内, 所以我们只保留 $x = 0$ 和 $x = 1$ 两点. 第二步告诉我们应该把 $x = -1$ 和 $x = 2$ 两点也加入列表.

有了一份全局最大值和最小值的可能候选者列表, 我们现在来到了第三步：求出它们所对应的函数值. 这很简单, 只需将它们逐一代入就可以了, 我们得到 $f(-1) = 44$, $f(0) = 1$, $f(1) = -12$ 及 $f(2) = 305$. 至于最后两步, 我们需要做的仅仅是从上述数值中选出最大的和最小的. 最大的是 305, 出现在 $x = 2$ 处, 所以 $x = 2$ 是该函数的全局最大值; 最小的是 -12, 出现在 $x =1$ 处, 所以 $x = 1$ 是该函数的全局最小值. 这样, 我们就找到了全局最大值和最小值.

在开始松懈之前, 让我们再更仔细地看一下刚才的函数 f. 首先, 注意到如果扩大它的定义域, 情况可能会由于两个原因而发生改变：一来端点会改变, 二来 $x = -2$ 处的临界点可能会进来搅局. 其次, 我们应该更仔细地看一下 $x = 0$ 处的临界点那里发生了什么. 它是局部最大值, 还是局部最小值, 又或者两者都不是? 对此一个方法是观察其图像. 如图 11-3 所示. 点 $(-1, 44)$ 比点 $(0, 1)$ 要高, 后者又比点 $(1, -12)$ 高. 所以在 $x = 0$ 处不可能有局部最大值, 也不可能有局部最小值. 但等等, 你可能会说 —— 图像也可能会像图 11-4 啊.

在图 11-4 中, $x = 0$ 是局部最大值. 但问题在于, 这样的话, 我们在 -1 和 0 之间引入了另一个局部最小值. 毕竟, 如果我们要求曲线连接 $(-1, 44)$ 和 $(0, 1)$, 同时又要求它在 $(0, 1)$ 处是个凸起, 那它势必曾降到比 1 还低的高度. 这意味着这里会

有一个低谷. 也就是说, 在 $x = -1$ 和 $x = 0$ 之间的某处有一个局部最小值! 但是, 这不可能, 因为在 $x = -1$ 和 $x = 0$ 之间没有临界点. 所以该函数的图像必定更接近于图 11-3, 进而结论是, x=0 既不是局部最大值也不是局部最小值.

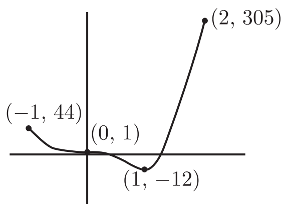


图　11-3

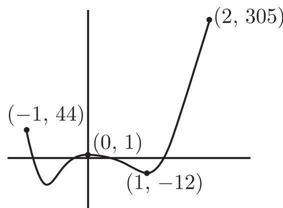


图　11-4

如果定义域不是有限的, 情况会变得略微复杂一些. 例如, 考虑下面两个函数 f 和 g, 它们的定义域都为 $[0, +\infty)$, 图像如图 11-5 所示.

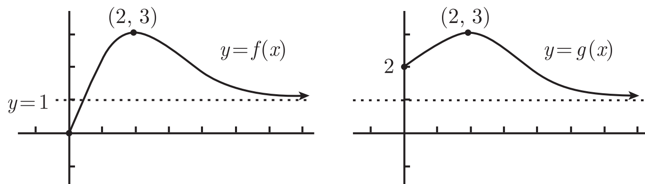


图　11-5

在两种情况下, 很显然 $x = 2$ 是临界点, 端点是 0 和 ∞. 但等一下, ∞ 并不是真正的端点, 因为它并不存在! 但无论如何, 让我们还是把它加入列表中, 因此列表包含 0, 2 和 ∞; 注意到对于 f 和 g 两个函数而言, 列表是一样的.

我们先来看函数 f. 可以看出 $f(0) = 0$, $f(2) = 3$, 而 $f(\infty)$ 只有当你考虑 $\lim\limits_{x\to\infty} f(x)$时才说得通. 该极限值为 1, 因为 $y = 1$ 为函数 f 的水平渐近线. 最大的函数值出现在 $x = 2$, 所以 $x = 2$ 是该函数的全局最大值. 最小的函数值出现在 $x = 0$, 所以 $x = 0$ 是该函数的全局最小值. 右 "端点" ∞ 甚至都没有出场.

那么函数 g 呢? 好吧, 这次 $g(0) = 2$, $g(2) = 3$, 而右端点由观察 $\lim\limits_{x\to\infty} g(x) = 1$ 可知. 最大的函数值依然出现在 $x = 2$, 所以 x =2 为函数的全局最大值. 但最小的函数值呢? 这个值当 $x \to \infty$ 时才能取到. 这是否意味着 ∞ 是全局最小值呢? 当然不是, 因为 ∞ 都不是一个数; 该函数 g 没有全局最小值.[①]

11.2　罗 尔 定 理

想象你正开车沿着高速公路行驶. 我看到你在一家加油站停了下来. 然后你继

① 另一方面, g 确实有一个全局下确界. 这一概念稍微超出了本书范围. 如果你想了解更多, 请参阅关于实分析的书.

续前行, 始终没有改变方向, 虽然你随时可以掉转方向. 过了一段时间, 我又在这家加油站看见了你, 但我不曾跟着你, 看你在这段时间里都做了什么. 我断定：在我不曾跟着看你的某个时刻, 你的车速为零.

为什么我会有信心这样说？其实, 有可能你从来就没有离开过加油站, 这样的话, 你的速度一直为零. 而如果你确实离开过加油站, 并往前开, 那你最终必定在某处掉了头, 否则你不可能又回到加油站. 那么当你停止前进开始掉头时会发生什么呢？你必定停下来过, 哪怕只是一瞬间! 你不可能掉转方向而不让车停下来. 这同我们在 6.4.1 节中研究过的上抛球运动的情况相似. 在球到达最高点的这一瞬间, 它的速度为零.

另一方面, 你还有可能曾离开加油站, 并倒着开. 在这种情况下, 你也必定曾在某个时刻将挡位由后退改为前进, 而结果是相同的：你在某处停下来过. 无论你向哪个方向走, 你都可能停下来过很多次; 但我知道你至少停下来过一次. 这就是罗尔定理所讲的内容.[①] 定理陈述如下.

> **罗尔定理** 假设函数 f 在闭区间 $[a,b]$ 内连续, 在开区间 (a,b) 内可导. 如果 $f(a)=f(b)$, 那么在开区间 (a,b) 内至少存在一点 c, 使得 $f'(c)=0$.

结合我们的例子, 设 $f(t)$ 是汽车在时刻 t 的位移. 这意味着 $f'(t)$ 是你在时刻 t 的速度. 时刻 a 和 b 是我在加油站看到你的时刻; $f(a)=f(b)$ 说明在时刻 a 和 b 你所在的位置相同 —— 都是在加油站. 最后, c 是你停下来的时刻, 因为 $f'(c)=0$. 罗尔定理告诉我你至少停下来过一次. 我不知道你是什么时候停下来的, 因为我没跟着你, 但我知道你肯定停下来过. (我假定你的车的运动是可导的, 这个假设在大多数情况下都很合理. 另一方面, 如果你从汽车碰撞测试假人的角度考虑, 或许车的运动在撞墙的那一瞬间不是可导的 $\cdots\cdots$)

现在, 让我们看一下罗尔定理适用的一些场合, 如图 11-6 所示.

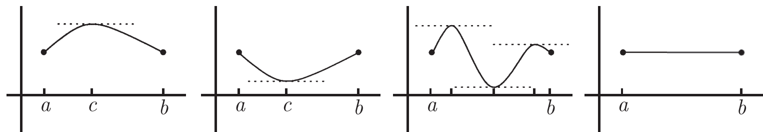


图 11-6

在前两个图中, 仅仅有一个可能的数值 c 使得 $f'(c)=0$. 在第三个图中, 有三个潜在的数值 c, 但这没有关系 —— 罗尔定理说的是至少有一个. 第四个图为常数函数图像, 导数始终为零. 这说明 c 可以是 a 和 b 之间的**任何值**. 接下来, 看一下罗尔定理不适用的一些场合, 如图 11-7 所示.

① 关于罗尔定理的证明请参见附录 A 的 A.6.7 节.

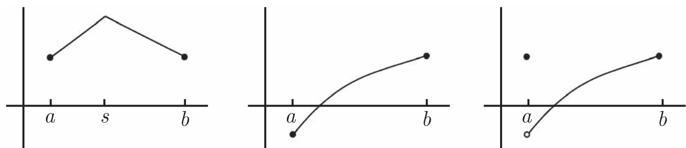


图　11-7

在上面三个图中, 导数都不会为零. 但这也没有关系, 因为罗尔定理在这三种情况下都不适用. 在第一个图中, 函数在开区间 (a,b) 内是不可导的, 因为在 s 点有一个尖点. 是的, 即使函数在一个点上不可导, 这也足以搞砸一切. 中间那个图, 函数是可导的, 但 $f(a) \neq f(b)$, 所以罗尔定理不适用. 在右边的图中, $f(a) = f(b)$ 且函数在开区间 (a,b) 内是可导的, 但该函数在闭区间 $[a,b]$ 内不是连续的: $x = a$ 这点让一切功亏一篑. 再一次地, 无法使用罗尔定理.

下面举一个罗尔定理应用的例子. 假设有一个函数 f 满足 $f'(x) > 0$ (对于所有的 x). 在 10.1.1 节中, 我们断言该函数一定满足水平线检验. 让我们用罗尔定理配合反证法证明这一点. 首先假设 f **不满足**水平线检验, 那么一定会有一条水平线, 比如说 $y = L$, 它与图像相交两次或更多. 假设这些交点中的两点的横坐标为 a 和 b, 则有 $f(a) = f(b) = L$. 由于 $f(a) = f(b)$, 所以可以用罗尔定理 (我们已经知道 f 是处处可导的, 所以它也一定是处处连续的). 这个定理指出, 在 a 和 b 之间一定存在一点 c 使得 $f'(c) = 0$. 但这是不可能的, 因为 $f'(x)$ 一直是正的! 所以该函数满足水平线检验.

现在来看一个更难一点的例子. 假设函数 f 的二阶导数处处存在且对于所有实数 $x, f''(x) > 0$. 这次的问题是, 证明函数与 x 轴至多有两个交点. 在开始解决问题之前, 让我们先稍微考虑一下这意味着什么? 你能想出一个函数, 对于所有实数 $x, f''(x) > 0$ 且与 x 轴没有交点吗? 那么一个交点呢? 两个交点呢? 如果你都能想得出来, 那再试试三个交点的情形. 不过不要在这上面浪费太多的时间, 因为这是不可能的. 的确, 我们的问题就是证明交点的个数不能超过两个.

事实上, 这里的关键在于: 如果有超过两个交点, 那就是说至少有三个交点. 让我们假设有两个以上的交点, 然后任意选取其中三个, 并这样分配记号 a, b 和 c 使得 $a < b < c$. 由于它们都为 x 轴截距, 所以有 $f(a) = f(b) = f(c) = 0$. 在闭区间 $[a,b]$ 内我们应用罗尔定理. 由于该函数在闭区间内处处连续, 开区间内处处可导, 所以一定有一点 p 在开区间 (a,b) 内使得 $f'(p) = 0$. 为什么我要用 p 呢? 因为这里 c 已经被占了!

接下来看闭区间 $[b,c]$. 再一次地, 由于 $f(b) = f(c)$, 根据罗尔定理, 在开区间 (b,c) 内一定存在一点 q 使得 $f'(q) = 0$. 别忘了, 我们已经有 $f'(p) = 0$. 啊哈, 现在可以在闭区间 $[p,q]$ 中使用罗尔定理, 但这次要用的不是函数 f, 而是其导函数 f'. 我们知道 $f'(p) = f'(q) = 0$, 所以根据罗尔定理, 在 (p,q) 区间内有一点 r 使得

$(f')'(r) = 0$. 等一下, $(f')'$ 就是二次导数 f''. 所以我们知道在开区间 (p,q) 内有一点 r 使得 $f''(r) = 0$. 这是个大问题, 因为我们已经假设对于所有实数 $x, f''(x) > 0$. 那么唯一的可能就是, 我们之前假设该函数与 x 轴有两个以上的交点是错误的. 因此交点个数不能超过两个, 问题解决.

顺便说一下, 刚才你想出来满足要求的函数 (对于所有实数 $x, f''(x) > 0$ 且与 x 轴分别有零、一或两个交点) 了吗? 如果没有, 试看一下 $f(x) = x^2 + C$, 其中 C 分别为正数、零或负数.

11.3 中值定理

想象你开始了另一段旅行, 我发现你在两个小时之内行驶了 100 英里. 因此, 你的平均速度为 50 英里/小时. 这并不是说你在整个行驶过程中速度始终维持在恰好 50 英里/小时. 现在, 我的问题是: 你的速度曾经达到过 50 英里/小时吗? 哪怕只是一瞬间.

答案是肯定的. 即使你在开始的第一个小时速度为 45 英里/小时, 第二小时为 55 英里/小时, 你仍然不得不从低速加速到高速. 而在这个过程中, 你有一瞬间的速度会是 50 英里/小时. 这是不可避免的! 不管你整个的行驶过程是怎样的, 如果你的平均速度为 50 英里/小时, 那么你会有至少一次瞬时速度为 50 英里/小时.[①] 当然, 你可能达到 50 英里/小时不止一次 —— 可能很多次, 甚至你能始终以 50 英里/小时的速度匀速行进. 这就引出了中值定理.

中值定理 假设函数 f 在闭区间 $[a,b]$ 内连续, 在开区间 (a,b) 内可导, 那么在开区间 (a,b) 内至少有一点 c 使得

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

这看起来有点儿古怪, 但实际上很说得通. 假设 $f(t)$ 是你在时刻 t 的位移, 你开始和结束的时刻分别为 a 和 b, 那么你的平均速度为多少? 你的位移为 $f(b) - f(a)$, 用的时间为 $b - a$, 所以上述等式的右边为你的平均速度. 另一方面, $f'(c)$ 是你在时刻 c 的瞬时速度. 中值定理指出, 在你的整个行程中至少会有一个时刻 c 使得你的瞬时速度等于平均速度.

让我们看这种情况的示意图. 假设函数图像如图 11-8 所示.

连接 $(a, f(a))$ 和 $(b, f(b))$ 两点的虚线斜率为 $\dfrac{f(b) - f(a)}{b - a}$. 根据中值定理, 某条切线的斜率与虚线斜率相同; 也就是说, 某条切线与虚线是平行的. 在上图中, 实际上有两条切线与虚线是平行的 —— 分别是在 c_0 和 c_1 处的切线. 任意其一都满足

① 再一次地, 所有这些都是基于一个很合理的假设 —— 你的汽车的运动是可导的!

定理中的 c.

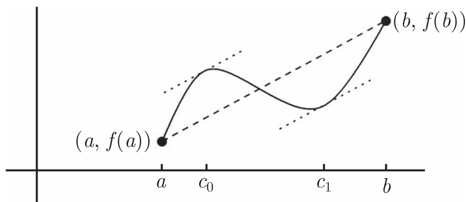


图　11-8

中值定理看上去很像罗尔定理. 实际上, 适用这两个定理的条件几乎是相同的. 在两个定理中, 函数 f 都要求在闭区间 $[a,b]$ 内连续, 在开区间 (a,b) 内可导. 但罗尔定理还要求 $f(a) = f(b)$, 中值定理则没要求这个. 实际上, 如果你对满足 $f(a) = f(b)$ 的函数 f 应用中值定理, 由于 $f(a) - f(b) = 0$, 于是你知道在开区间 (a,b) 内有一点 c 使得 $f'(c) = 0$. 所以中值定理可以推导出罗尔定理.

下面来看一些如何应用这个定理的例子. 首先, 如何证明方程

$$2x\mathrm{e}^{x^2} - \mathrm{e} + 1 = 0$$

有解？一个方法是使用介值定理 (参见 5.1.4 节), 你可以现在试试. 不过在这里, 我建议对区间 $[0,1]$ 上的函数 $f(x) = \mathrm{e}^{x^2}$ 使用中值定理. 这是可行的, 因为该函数在其定义域内是处处连续并可导的. 根据中值定理, 在闭区间 $[0, 1]$ 内至少存在一点 c 满足

$$f'(c) = \frac{f(1) - f(0)}{1 - 0}.$$

显然需要求出 $f'(x)$. 使用链式求导法则, 应该可以证明 $f'(x) = 2x\mathrm{e}^{x^2}$. 所以上述方程变为

$$2c\mathrm{e}^{c^2} = \frac{\mathrm{e}^{1^2} - \mathrm{e}^{0^2}}{1 - 0} = \mathrm{e} - 1.$$

这样就得到 $2c\mathrm{e}^{c^2} - \mathrm{e} + 1 = 0$, 从而证明了原始方程有解. 事实上, 这也证明了在 0 与 1 之间存在一个解.

下面是个难点的例子. 假设有这样一个函数, 对于所有的实数 x 处处可导并且 $f'(x) > 4$. 问题是, 如何证明这个函数 $y = f(x)$ 的图像与线性函数 $y = 3x - 2$ 最多只有一个交点. 先试一下, 看你是否可以在继续阅读之前解决它.

好吧, 究竟该怎样解决这个问题呢？事实上, 这同上一节中的罗尔定理的例子很相似. 首先, 注意到如果点 (x, y) 是同时满足函数 $y = f(x)$ 和线性函数 $y = 3x-2$ 的点, 那么一定会有 $f(x) = 3x - 2$. 这个方程对于绝大多数 x 并不成立! 它只对交点的 x 成立. 依然用反证法, 假设交点不止一个. 任意选取其中两个, 并这样分配记号 a 和 b, 使得 $a < b$. 由于它们是交点, 我们知道有 $f(a) = 3a - 2$ 和 $f(b) = 3b - 2$.

又由于该函数对于所有实数都处处可导且连续, 根据中值定理, 在开区间 (a,b) 内一定有一点 c 使得

$$f'(c)=\frac{f(b)-f(a)}{b-a};$$

代入 $f(a)=3a-2$ 和 $f(b)=3b-2$, 可得

$$f'(c)=\frac{(3b-2)-(3a-2)}{b-a}=\frac{3\,(b-a)}{b-a}=3.$$

但这是不可能的, 因为对于所有 x, $f'(x)>4$. 因此, 最多只能有一个交点.

这样就完成了这个证明, 但你可能会想知道对于这个问题的另一种解读. 想象一辆车 A 正以 3 英里/小时的速度前进, 它的开始位移为 -2, 那么它在任意时刻 t 的位移表达式为 $3t-2$. 假设你在任意时刻 t 的位移是 $f(t)$, 那么 $f'(t)>4$ 意味着你在任意时刻的速度永远大于 4 英里/小时 (与 A 车同方向). 所以问题变为, 证明你不能与 A 车相遇的次数超过一次. 假设你们相遇超过一次, 那么由于 A 车的速度恒为 3 英里/小时, 你至少在某一时刻的速度为 3 英里/小时. 但这是不可能的, 因为你的速度一直都大于 4. 如果这样想的话, 这个问题的证明就很说得通了!

中值定理的几个推论

长久以来, 一些关于导数的结论我们一直不加证明就直接使用. 比如说, 如果一个函数的导数始终为零, 那么这个函数一定为常数函数. 诸如这样的事实看上去显而易见, 但其实是需要证明的. 下面就让我们用中值定理去证明三个关于导数的有用事实.

(1) 假设函数 f 在开区间 (a,b) 内的任意一点的导数都为零. 这意味着该函数的图像是水平的. 事实上, 很显然该函数在这个区间内是常数函数. 但怎样证明呢? 首先, 在该区间内固定一点 S, 然后在该区间内任取一点 x(x 不同于 S). 根据中值定理, 在 x 和 S 之间一定存在一点 c 满足

$$f'(c)=\frac{f(x)-f(S)}{x-S}.$$

由于我们已经假设函数的导数始终为零, 这说明 $f'(c)$ 也一定为零. 所以上述方程变为

$$f'(c)=\frac{f(x)-f(S)}{x-S}=0,$$

这意味着 $f(x)=f(S)$. 如果设 $C=f(S)$, 那么对于所有在该区间内的 x 有 $f(x)=C$, 所以该函数为常数函数. 于是我们有这样的结论:

> 如果对于在定义域 (a,b) 内的所有 x, 都有 $f'(x)=0$, 那么函数 f 在开区间 (a,b) 内为常数函数.

事实上, 在 10.2.2 节中已经使用过这个结论. 在那里, 如果 $f(x) = \sin^{-1}(x) + \cos^{-1}(x)$, 那么对于开区间 $(-1,1)$ 内的所有 x, $f'(x) = 0$. 于是我们得出结论, 函数 f 在该区间内为常数函数. 又由于 $f(0) = \pi/2$, 实际上得到：对于所有在开区间 $(-1,1)$ 内的 x 都有 $\sin^{-1}(x) + \cos^{-1}(x) = \pi/2$.

(2) 假设两个可导函数有相同的导数. 那么它们是同一个函数吗？不一定, 它们可能相差一个常数. 例如, 函数 $f(x) = x^2$ 和 $g(x) = x^2 + 1$ 有相同的导数 $2x$, 但很明显, 这两个函数是不同的函数. 那么还有其他方法使这两个函数处处有相同的导数吗？答案是否定的, 相差一个常数是唯一的方法.

如果对于任意实数 x 都有 $f'(x)=g'(x)$, 那么有 $f(x)=g(x)+C$ (C 为常数).

事实上, 使用前面的事实 (1) 可以很容易证明这一点. 假设对于所有 x, $f'(x) = g'(x)$. 现在令 $h(x) = f(x) - g(x)$. 对等式两边同时求导, 有 $h'(x) = f'(x) - g'(x) = 0$, 所以 h 为常数函数. 也就是说, $h(x) = C$ (C 为某个常数). 这意味着 $f(x) - g(x) = C$ 或 $f(x) = g(x) + C$. 函数 f 和 g 确实只相差一个常数. 这个事实对于我们后面章节的积分学习将是非常有用的.

(3) 如果函数 f 的导函数始终为正, 那么该函数为**增函数**. 也就是说, 如果 $a < b$, 则有 $f(a) < f(b)$. 换句话说, 在图像上任取两点, 那么左边的点一定低于右边的. 当你从左向右看时, 此曲线一点点变高. 但为什么会这样呢？假设对于所有 x, 有 $f'(x) > 0$ 并且假设 $a < b$. 根据中值定理, 在开区间 (a,b) 内至少存在一个常数 c 使得

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

这意味着 $f(b) - f(a) = f'(c)(b - a)$. 由于 $f'(c) > 0$ 且 $b - a > 0$, 所以等式的右边为正. 这样我们有 $f(b) - f(a) > 0$, 因此 $f(b) > f(a)$, 所以该函数的确为增函数. 另一方面, 如果对于所有 x, $f'(x) < 0$, 那么这样的函数是**减函数**; 也就是说, 如果 $a < b$, 则有 $f(a) > f(b)$. 证明的方法是基本一样的.

11.4　二阶导数和图像

到目前为止, 还没有太多讨论过二阶导数. 我们只用它来定义过加速度, 仅此而已. 但实际上, 二阶导数能告诉你很多关于函数图像的信息. 例如, 假设知道对于开区间 (a,b) 内的所有 x, $f''(x) > 0$. 而如果把二阶导数看作导数的导数, 那么可以把二阶导数写为 $(f')'(x) > 0$. 这意味着导函数 $f'(x)$ 始终是增函数.

那又怎么样呢？好吧, 如果知道导函数为增函数, 这意味着函数图像会变得越来越 “陡峭”, 如图 11-9 所示. 在紧靠 $x = a$ 的右边, 登山者轻松惬意：斜率为负.

但情况逐渐变得越来越艰难. 山势变得越来越平坦, 直到完全水平的 $x = c$ 处, 然后随着斜率逐渐增加, 山势变得越来越陡峭, 直到 $x = b$ 处. 这里的要点在于, 从 $x = a$ 到 $x = b$ 斜率始终在增加. 而这也正是式子 $f''(x) > 0$ 所暗示的.

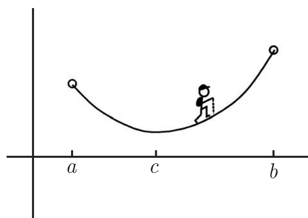


图 11-9

我们需要用某种方式描述这样的行为. 如果函数的斜率在某段区间 (a,b) 内为增函数, 或换言之, 它的二阶导数在该区间内始终为正 (假设二次导数存在), 那么我们说该函数在该区间是凹向上的. 图 11-10 是一些凹向上函数的图像.

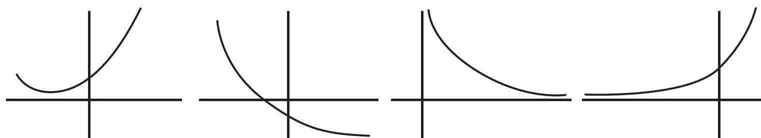


图 11-10

它们看上去都像碗的一部分. 注意到仅仅通过 $f''(x) > 0$ 我们无法判断一阶导数 $f'(x)$ 的正负. 确实, 上述图像的中间两个的一阶导数为负, 最右边的一阶导数为正, 最左边的一阶导数由负到正.

如果二阶导数 $f''(x)$ 为负, 那么情况就反了过来. 它们看上去就像倒扣的碗. 如果在某段区间内函数 f 的二阶导数始终为负, 那么就称 f 在该区间是凹向下的.[1] 图 11-11 是一些凹向下函数的图像.

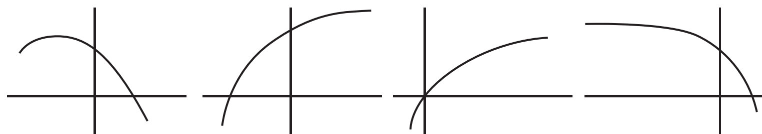


图 11-11

在这些图像中, 函数的导函数都为减函数. 这意味着你会发现在这座山上行进越来越容易：如果你是在上山, 山势会越来越平坦; 而如果你是在下山, 山势则会越来越陡峭 (你都是从左往右).

① 如果你记不清哪一个是凹向上, 哪一个是凹向下, 那么下面这两个尾韵也许能帮助你："Like a cup, con cave up; Like a frown, concave down." (茶杯样, 凹向上; 皱眉相, 凹向下.)

当然, 凹性并不需要每一个地方都一样: 它可以改变. 如图 11-12 所示, 在 $x = c$ 点的左边, 图像是凹向下的; 而在 $x = c$ 点的右边, 图像是凹向上的. 这时, 我们称 c 点为函数的**拐点**, 因为函数在 c 点改变了它的凹性.

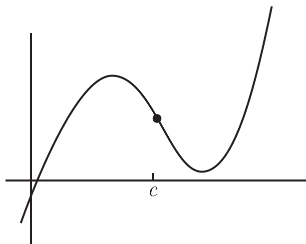


图　11-12

关于拐点的更多说明

在图 11-12 中, 我们看到在 c 点的左边二阶导数小于零, 在 c 点的右边二阶导数大于零. 那么在 c 点的二阶导数又是怎样的呢? 它肯定为零, 因为所有的一切都是连续平滑的. 一般而言, 如果 c 点为拐点, 那么 $x = c$ 点两侧的二次导数的符号必定是相反的, 前提是当 x 接近于 c 点时 $f''(x)$ 确实存在. 在那种情况下, 必有以下结论:

如果 $x = c$ 点是函数 f 的拐点, 则有 $f''(c) = 0$.

但另一方面, 如果 $f''(c) = 0$, 则 c 点可能是也可能不是拐点! 也就是说,

如果 $f''(c) = 0$, 则 c 点不一定都是函数 f 的拐点.

例如, 假设函数 $f(x) = x^4$, 那么 $f'(x) = 4x^3$, $f''(x) = 12x^2$. 在 $x = 0$ 点, 它的二阶导数为零, 因为 $f''(0) = 12(0)^2 = 0$. 那么 $x = 0$ 是拐点吗? 答案是否定的. 图 11-13 为该函数图像.

从图像中可以看出, 函数在其定义域内都是凹向上的, 所以在 $x = 0$ 这点该函数并没有改变它的凹性. 也就是说, 尽管 $f''(0) = 0$, $x = 0$ 这点并不是它的拐点.

另一方面, 如果你想找拐点, 确实应该找二阶导数为零的点. 这样做至少可以缩小寻找范围, 然后我们可以再逐一检验. 例如, 假设 $f(x) = \sin(x)$, 那么 $f'(x) = \cos(x)$, $f''(x) = -\sin(x)$. 当 x 的值为 π 的整数倍时, 该函数的二阶导数为零. 此时, $f''(0) = -\sin(0) = 0$, 那么 $x = 0$ 是拐点吗? 让我们看一下其图像, 如图 11-14 所示. 是的, $x = 0$ 是拐点: $\sin(x)$ 在 0 的左边是凹向上, 而在 0 的右边是凹向下. 注意到在 $x = 0$ 处的切线穿过曲线 $y = \sin(x)$. 这对拐点而言是典型的: 在拐点一边曲线必定在切线之上, 而在另一边在切线之下.

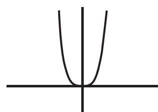


图 11-13

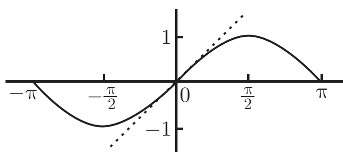


图 11-14

11.5 对导数为零点的分类

现在是时候把前面的部分理论应用到实际问题中去了. 假设有一个函数 f 以及数 c 使得 $f'(c) = 0$. 除了可以确定地说 c 点是函数 f 的临界点, 你还可以说出什么? 事实证明, 这里仅有三种常见可能性: $x = c$ 可能为局部最大值; 也可能为局部最小值; 还可能为水平拐点, 也就是说, 这点不仅是拐点, 通过该点的切线也是水平的.[①] (也有可能对于所有接近于 c 的 x, $f(x)$ 是常数函数, 但这样的话, c 就既是局部最大值又是局部最小值.) 不论如何, 图 11-15 是这几种常见可能性的示意图.

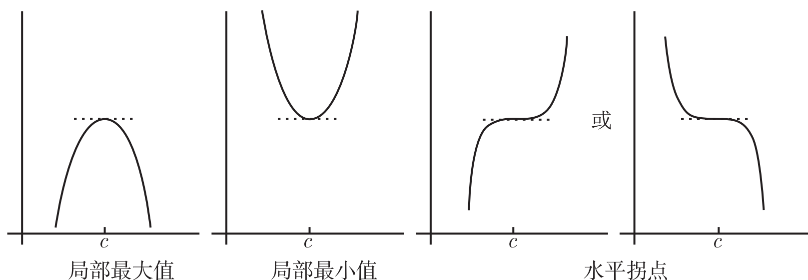


图 11-15

在每一种情况下, 切线都是水平的: 这是你只知道 $f'(c) = 0$ 时所能得出的唯一结论. 那么怎样才能判断究竟属于上图中的哪种情况呢? 有两个方法, 一个只需用到一阶导数, 另一个则需用到二阶导数. 当使用一阶导数时, 需要观察在 $x = c$ 附近的一阶导数的符号 (是正还是负). 另一方面, 当使用二阶导数时, 需要考虑在 $x = c$ 点的二阶导数的符号. 让我们逐一来看这两个方法.

11.5.1 使用一阶导数

让我们再看一下上边的示意图, 但这次在 $x = c$ 点两侧画一些切线, 如图 11-16 所示.

在第一个中, $x = c$ 点为局部最大值. 在 c 点的左侧, 图像的斜率为正: 也就是说, 在那一部分定义域函数为增函数 (参见 11.3.1 节). 另一方面, 在 c 点的右侧, 图

① 另一种可能性是在临界点附近的凹性甚至不是良定义的. 例如 $f(x) = x^4 \sin(1/x)$ 这个函数, 当 x 趋于临界点 0 时, 二阶导数的符号反复振荡, 所以凹性也在不停变化!

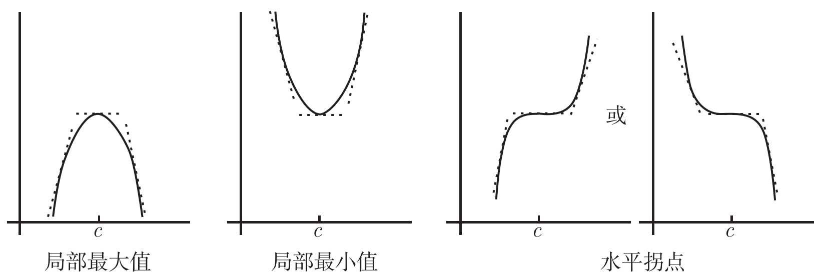


图　11-16

像的斜率为负; 也就是说, 在那一部分定义域函数为减函数. 很显然, 如果随着你从左往右, 斜率由正变负, 那么斜率为零的那一点必定是局部最大值.

对于第二个, 情况恰恰相反. 如果从左往右, 斜率由负变正, 那么斜率为零的那一点必定是局部最小值. 在第三个中, 除 c 点外, 斜率始终为正; 在第四个中, 除 c 点外, 斜率始终为负. 后两者中的 c 点均为拐点：点两侧的导数的斜率并没有改变符号.



下面是对上述观察的总结. 假设 $f'(c) = 0$, 这时有：

- 如果从左往右通过 c 点, $f'(x)$ 的符号由正变负, 那么 c 点为局部最大值;
- 如果从左往右通过 c 点, $f'(x)$ 的符号由负变正, 那么 c 点为局部最小值;
- 如果从左往右通过 c 点, $f'(x)$ 的符号不发生变化, 那么 c 点为水平拐点.



例如, 如果函数 $f(x) = x^3$, 那么我们有 $f'(x) = 3x^2$. 由于当 $x = 0$ 时导数为零, 所以 $x = 0$ 一定是局部最大值、局部最小值或水平拐点中的一种. 但到底是哪一种呢？由于当 $x \neq 0$ 时, 导函数始终为正, 则从左往右通过 $x = 0$ 时, 导数的符号不发生变化, 所以该点一定为拐点. 你可以画函数图像检验一下! (在 11.5.2 节你会看到该函数图像.)



下面是另一个例子. 如果设 $f(x) = x\ln(x)$, 那么函数 f 的局部最大值、局部最小值和水平拐点又在哪里呢？首先, 可以使用乘积法则求得 $f'(x) = \ln(x) + 1$. (自己检验一下!) 接下去需要求解方程 $f'(x)=0$, 即 $f'(x)=\ln(x)+1=0$.

通过重新整理, 我们得到 $\ln(x) = -1$, 两边同时取幂, 得到 $x = \mathrm{e}^{-1} = 1/\mathrm{e}$. 这是唯一的候选者, 但它是哪种类型的临界点呢？

好吧, 让我们看一下 $f'(x) = \ln(x) + 1$ 在 x 接近 $1/\mathrm{e}$ 时的符号. 最简单的方式是画出导函数 $y = f'(x)$ 的图像草图. 我们所需做的只是把 $\ln(x)$ 的图像向上平移一个单位, 如图 11-17 所示. 从图像中可以看出, 随着从左往右通过 $x = 1/\mathrm{e}$ 导函数由负变正, 所以 $x = 1/\mathrm{e}$ 必定为局部最小值. 那么在该点的函数值又是多少呢？把 $x = 1/\mathrm{e}$ 代入原函数, 得到 $f(1/\mathrm{e}) = (1/\mathrm{e})\ln(1/\mathrm{e}) = -1/\mathrm{e}$, 因为其中 $\ln(1/\mathrm{e}) = \ln(\mathrm{e}^{-1}) = -\ln(\mathrm{e}) = -1$. 因此, 该函数在点 $(1/\mathrm{e}, -1/\mathrm{e})$ 有局部最小值. 它在那个局部的图像应该如图 11-18 所示. 但正如你所看到的, 我们还不知道其他部

分的图像如何. 我们将在 12.3.2 节将它补完.

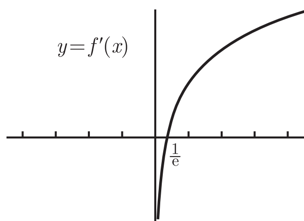


图 11-17

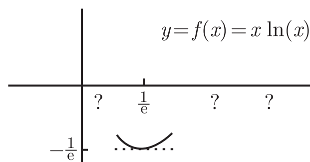


图 11-18

11.5.2 使用二阶导数

再来看一下当 $f'(c)=0$ 时几种常见可能性, 如图 11-19 所示.

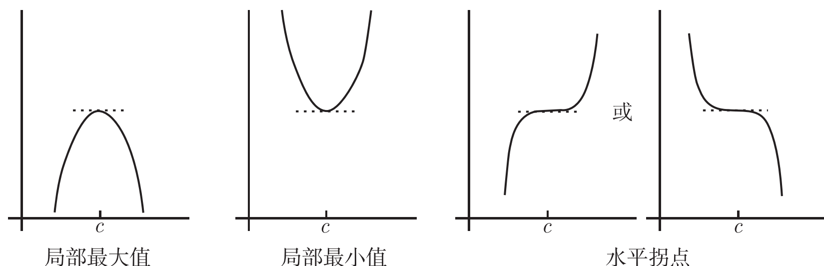


图 11-19

假设 $f''(c)>0$. 从 11.4 节可知, 这样的函数 $y=f(x)$ 的图像在 $x=c$ 附近是凹向上的. 上面只有第二个满足条件, 这时 $x=c$ 是局部最小值. 类似地, 如果 $f''(c)<0$, 那么图像就是凹向下的, 也就是上面第一个的情形, 此时 $x=c$ 为局部最大值.

这个方法相当管用, 但它也有一个缺陷：如果 $f''(c)=0$, 那么可能遇到上述四种情况的任意一种! 例如, 假设 $f(x)=x^3, g(x)=x^4$. 我们有 $f'(x)=3x^2$, 所以 $f'(0)=0$. 接下来让我们用它的二阶导数去对这个临界点进行分类. 由于 $f''(x)=6x$, 则有 $f''(0)=0$.

另一方面, 函数 g 呢? 在 11.4.1 节中已经求得 $g'(x)=4x^3$, 所以 $g'(0)=0$. 这里的 $x=0$ 又是什么类型的临界点呢? 让我们用二阶导数来检验一下：$g''(x)=12x^2$, 所以 $g''(0)=0$.

在这两种情况下, 在临界点 $x=0$ 的二阶导数都为零. 而从图 11-20 可以看出, 函数 f 在 $x=0$ 有一个拐点, 函数 g 在 $x=0$ 则有一个局部最小值.

在这样的情况下, 使用二阶导数并没有什么用处. 当二阶导数为零时, 你无异于两眼一抹黑, 完全无法分辨自己面对的究竟是局部最大值、局部最小值还是水平拐点. 下面是一些总结. 假设 $f'(c)=0$, 则有：



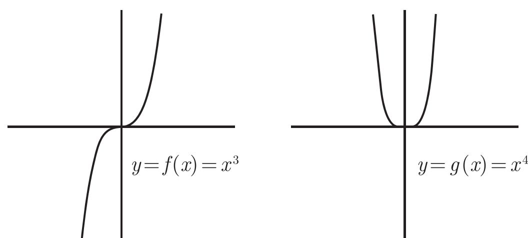


图　11-20

- 如果 $f''(c) < 0$, 那么 $x = c$ 为局部最大值;
- 如果 $f''(c) > 0$, 那么 $x = c$ 为局部最小值;
- 如果 $f''(c) = 0$, 那么你无法判断发生了什么! 需要使用上一节讲过的一阶导数方法.

是的, 一阶导数方法更好, 虽然它略微复杂一点. 它在任何情况下都可以使用, 而不像二阶导数方法那样有局限性. 下面是一个两种方法都适用的例子. 假设函数 $f(x) = x\ln(x)$. 这是上一节中使用过的例子! 我们已经通过一阶导数方法发现 1/e 是该函数的局部最小值. 现在就用二阶导数方法再试一下.

首先, 忆及 $f'(x) = \ln(x)+1$, 所以 $f'(1/\mathrm{e}) = 0$. 很容易看出, $f''(x) = 1/x$. 所以当 $x = 1/\mathrm{e}$ 时, 有 $f''(1/\mathrm{e}) = \mathrm{e}$, 而这是个大于零的数. 所以该函数在 $x = 1/\mathrm{e}$ 的凹性为凹向上, 就像一个碗的形状; 而根据前面的总结, $x = 1/\mathrm{e}$ 的确为局部最小值.

第 12 章　绘制函数图像

现在是时候来看一下绘制函数 $y = f(x)$ 的图像的一般方法了. 当我们绘制函数图像时, 并没有打算追求完美, 而只是希望能把函数的主要特征表现出来. 确实, 我们将用到已经掌握的微积分知识：用极限去找渐近线, 用一阶导数去找极大值和极小值, 用二阶导数去找函数的凹性. 以下是我们将要讨论的话题：

- 建立符号表格这种有用技巧;
- 绘制函数图像的一般方法;
- 如何应用该方法的五个例子.

12.1　建立符号表格

假设要绘制函数 $y = f(x)$ 的图像. 对于任意的 x 值, 它所对应的 $f(x)$ 可能为正, 可能为负, 可能为零, 也可能在该点没有定义. 幸运的是, 如果该函数除了少量点外都是连续的, 并且你能找到它所有的零点以及不连续点, 那么通过使用符号表格就可以很容易地看出 $f(x)$ 在哪里为正, 哪里为负了.

下面就是具体做法：首先, 以递增的顺序列出所有零点和不连续点. 例如, 如果

$$f(x) = \frac{(x-3)(x-1)^2}{x^3(x+2)},$$

那么零点分别为 3 和 1, 不连续点分别在 0 和 -2. 按递增的顺序排列就是 -2, 0, 1, 3. 现在, 建立一个三行多列的表格, 前两行分别标为 x 和 $f(x)$, 第三行暂时留白. 接下来, 把刚才列出的零点以及不连续点填入表格的第一行, 并且每个数的左右都要留出空格, 如图 12-1 所示.

x		-2		0		1		3	
$f(x)$									

图　12-1

现在你可以填充第二行的部分空格 —— 函数值为 0 的点直接填 0, 不连续点填以星号, 见图 12-2.

x		-2		0		1		3	
$f(x)$		$\star$		$\star$		0		0	

图　12-2

接下来, 在第一行的每两个数之间及其前后选取你喜欢的数. 在我们的例子中, 可能会为 -2 的左边选 -3, 在 -2 和 0 之间选 -1, 如此等等. 最后表格看上去会如图 12-3.

x	-3	-2	-1	0	$\frac{1}{2}$	1	2	3	4
$f(x)$		$\star$		$\star$		0		0	

图 12-3

我们也可以选 -4 而不是 -3, 选 1/3 而不是 1/2 —— 这是无关紧要的. 我们可以选取两个数之间的**任意数**. 接下去要做的是, 判断所选的数所对应的函数值的正负. 例如当 $x=-3$ 时,

$$f(-3)=\frac{(-3-3)(-3-1)^2}{(-3)^3(-3+2)}=-\frac{32}{9}.$$

所以我们可以在 -3 的下面填一个减号. 实际上没有必要完全求出函数值, 因为我们不怎么关心 $f(-3)$ 的具体值, 只关心它的正负. 我们只需通过判断每一个因式的正负去判断整个算式的正负. 具体说, 当 $x=-3$ 时, $(x-3)$ 为负, $(x-1)^2$ 为正, (必然为正, 因为这是个平方表达式!) x^3 为负, $(x+2)$ 也为负. 这样, 总的效果是

$$\frac{(-)(+)}{(-)(-)}=-,$$

所以 $f(-3)$ 为负. 现在试着对其他数作同样的分析, 应该得到图 12-4.

x	-3	-2	-1	0	$\frac{1}{2}$	1	2	3	4
$f(x)$	$-$	$\star$	$+$	$\star$	$-$	0	$-$	0	$+$

图 12-4

这里的关键不是 $f(-3)$ 为负, 而是 $f(x)$ 对于所有的 $x<-2$ 都为负. 数 -3 仅仅是 $(-\infty,-2)$ 中所有数的一个代表性样本. $f(-3)$ 的正负体现了该函数在 $(-\infty,-2)$ 区间内的正负. 类似地, 由于 $f(-1)$ 是正的, $f(x)$ 在 $(-2,0)$ 的整个区间内是正的. 这样的表格已然告诉了我们关于函数 $y=f(x)$ 的很多信息, 对此将在 12.3.1 节再展开讨论.



下面是另一个例子. 假设函数

$$f(x)=x^2(x-5)^3.$$

我们在 10.1.4 节中已经见过这个函数了. 现在用符号表格再来更仔细看一

x	-1	0	2	5	6
$f(x)$	$-$	0	$-$	0	$+$

图　12-5

下这个函数. 该函数的零点只有 $x=0$ 和 $x=5$, 但没有不连续点. 所以特殊点为 0 和 5. 接下来需要填表. 在 0 的左边我选 -1, 在 0 和 5 之间我选 2, 在 5 的右边我选 6. 所以我们的表格大致如图 12-5.

下面是我如何得到在 -1, 2 和 6 处的符号.

- 当 $x=-1$ 时, x 和 $(x-5)$ 都为负. 因此, $f(-1)$ 为 $(-)^2(-)^3=(+)(-)=(-)$.
- 当 $x=2$ 时, x 为正, $(x-5)$ 为负. 因此, $f(2)$ 为 $(+)^2(-)^3$, 仍然为负.
- 当 $x=6$ 时, x 和 $(x-5)$ 都为正. 因此, $f(6)$ 为 $(+)^2(+)^3=(+)$.

我们会在 12.3.3 节绘制该函数图像时再次回到这个表格. 不过现在, 先看看如何建立一阶导数和二阶导数的符号表格.

12.1.1　建立一阶导数的符号表格

正如 11.3.1 节所述, 一阶导数的符号可以告诉我们关于函数的很多信息. 导数为正, 函数为增函数; 导数为负, 函数为减函数; 导数为 0, 函数有局部最大值、局部最小值或水平拐点. 一个一阶导数的符号表格能把所有这些信息简明扼要地总结出来.

方法同刚才 $f(x)$ 的符号表格所用方法是一样的, 只是现在你是基于 $f'(x)$. 另外一个不同之处是, 当 $f'(x)$ 为 0 时, 我们在第三行画一条小水平横线; 当 $f'(x)$ 大于 0 时, 我们画一条斜向上的斜线; 当 $f'(x)$ 小于 0 时, 我们画一条斜向下的斜线.



让我们看看它如何应用于刚才的例子 $f(x)=x^2(x-5)^3$. 在 10.1.4 节中已经算得 $f'(x)=5x(x-5)^2(x-2)$. (如果你不想翻回去看, 可以自己重新计算一下!) 这意味着当 $x=0$, $x=2$ 或 $x=5$ 时 $f'(x)=0$. 让我们选取它们之间的一些点：在 0 的左边选 -1; 在 0 和 2 之间选 1; 在 2 和 5 之间选 3; 最后, 在 5 的右边选 6. 我们的符号表格目前看上去如图 12-6.

x	-1	0	1	2	3	5	6
$f'(x)$		0		0		0	

图　12-6



接下来, 需要判断在所选取的这些新点上 $f'(x)$ 的符号. 例如, 当 $x=-1$ 时, $5x$ 为负, $(x-5)$ 为负, $(x-2)$ 也为负, 所以 $f'(-1)$ 的符号为 $(-)(-)^2(-)=(+)$. 我留给你重复这个练习, 判断其他几个点上的符号, 确保得到图 12-7.

x	-1	0	1	2	3	5	6
$f'(x)$	+	0	−	0	+	0	+
	/	—	\	_	/	—	/

图 12-7

注意我在第三行是怎样画线的：当 $f'(x)$ 为正时, 画斜向上的线; 当 $f'(x)$ 为负时, 画斜向下的线; 当 $f'(x)$ 为 0 时, 画水平的线. 这样我们马上就知道, 当 $x<0$ 和 $x>2$ 时 f 为增函数; 当 $0<x<2$ 时, f 为减函数. 上述表格也告诉我们, $x=0$ 为局部最大值, $x=2$ 为局部最小值, $x=5$ 为水平拐点. 我们会在 12.3.3 节绘制该函数图像时再次用到这个表格.

一点提醒：表格第三行中的短线旨在作为你作图时的导引. 函数图像很有可能根本不像把这些短线连起来后的样子! 所以应该只用那一行中的信息来理解函数在哪里是增函数、在哪里是减函数, 或者在哪里暂时是水平的.

12.1.2 建立二阶导数的符号表格

我们也已经看到二阶导数的重要性 (回顾一下 11.4 节). 当二阶导数为正时, 函数图像是凹向上的; 为负时, 图像是凹向下的; 为 0 时, 你可能得到也可能得不到一个拐点. 一张二阶导数的符号表格会告诉我们这些信息.

方法同函数值或一阶导数的符号表格所用方法是一样的, 只是现在第三行要用来表示函数图像是凹向上还是凹向下. 当 $f''(x)$ 为正时, 画一个开口向上的小抛物线; 为负时, 画一个开口向下的; 为 0 时, 画一个点.

回到刚才的例子 $f(x)=x^2(x-5)^3$, 我们已经知道 $f'(x)=5x(x-5)^2(x-2)$. 为了对这再求导, 需将 x 和 $(x-2)$ 合并在一起, 得到 $f'(x)=5(x-5)^2(x^2-2x)$. 接下来, 可以应用乘积法则, 得到

$$f''(x)=5((x^2-2x)\times(2(x-5))+(x-5)^2(2x-2)).$$

提出公因式 $(x-5)$ 并重新整理, 得到 $f''(x)=10(x-5)(2x^2-8x+5)$. 实际上, 可以使用二次方程求根公式去求 $2x^2-8x+5=0$ 的解, 求得解为 $2\pm\dfrac{1}{2}\sqrt{6}$. 所以可以把 $f''(x)$ 彻底地因式分解为

$$f''(x)=20\left(x-\left(2-\frac{1}{2}\sqrt{6}\right)\right)\left(x-\left(2+\frac{1}{2}\sqrt{6}\right)\right)(x-5).$$

这意味着当 $x=2-\dfrac{1}{2}\sqrt{6}$, $x=2+\dfrac{1}{2}\sqrt{6}$ 和 $x=5$ 时, $f''(x)$ 的值为 0. 这样我们初步得到 $f''(x)$ 的表格如图 12-8.

x		$2-\frac{1}{2}\sqrt{6}$		$2+\frac{1}{2}\sqrt{6}$		5	
$f''(x)$		0		0		0	

图 12-8

现在, 要填充空白处. 如果能知道 $2 \pm \frac{1}{2}\sqrt{6}$的大概值会很有帮助, 所以让我们试着不用计算器去估算一下. 你看, $\sqrt{6}$ 是在 2 和 3 之间 (因为 6 是在 4 和 9 之间), 所以 $\frac{1}{2}\sqrt{6}$ 是在 1 和 3/2 之间. 这意味着 $2-\frac{1}{2}\sqrt{6}$ 是在 $2-\frac{3}{2}=\frac{1}{2}$ 和 $2-1=1$ 之间, 而 $2+\frac{1}{2}\sqrt{6}$ 是在 2+1=3 和 $2+\frac{3}{2}=3\frac{1}{2}$ 之间. 所以我们在 $2-\frac{1}{2}\sqrt{6}$ 的左边选 0; 在 $2-\frac{1}{2}\sqrt{6}$ 和 $2+\frac{1}{2}\sqrt{6}$ 之间选 2; 在 $2+\frac{1}{2}\sqrt{6}$ 和 5 之间选 4; 最后, 在 5 的右边选 6. 这样就会得到图 12-9.

x	0	$2-\frac{1}{2}\sqrt{6}$	2	$2+\frac{1}{2}\sqrt{6}$	4	5	6
$f''(x)$	$-$	0	$+$	0	$-$	0	$+$
	$\frown$	$\cdot$	$\smile$	$\cdot$	$\frown$	$\cdot$	$\smile$

图 12-9

确保你理解了上表中我所填写的所有符号是正确的. 例如当 $x=0$ 时, $f''(x)$ 的三个因式都是负的, 所以乘积也是负的. 还要注意到我如何在第三行画小抛物线. 你可以很清楚地看到, 当 $2-\frac{1}{2}\sqrt{6}<x<2+\frac{1}{2}\sqrt{6}$ 或 $x>5$ 时, 图像是凹向上的; 而当 $x<2-\frac{1}{2}\sqrt{6}$ 或 $2+\frac{1}{2}\sqrt{6}<x<5$ 时, 图像是凹向下的. 同时, 点 $2-\frac{1}{2}\sqrt{6}$, $2+\frac{1}{2}\sqrt{6}$和 5 都是拐点, 因为在这些点的左右两侧的凹性正好相反. 我们会在 12.3.3 节再回到这个表格.

再看另一个例子. 假设 $g(x)=x^9-9x^8$. 很容易算得 $g'(x)=9x^8-72x^7$, $g''(x)=72x^7-72\times 7x^6=72x^6(x-7)$. 所以当 $x=0$ 或 $x=7$ 时 $g''(x)=0$. 让我们选 $x=-1$, $x=3$ 和 $x=8$ 作为填充的点. 我留给你来证明 $g''(-1)<0$, $g''(3)<0$ 和 $g''(8)>0$. 最终 $g''(x)$ 的符号表格应该大致如图 12-10.

x	-1	0	3	7	8
$g''(x)$	$-$	0	$-$	0	$+$
	$\frown$	$\cdot$	$\frown$	$\cdot$	$\smile$

图 12-10

可以发现, $x=0$ 并不是拐点, 因为在 $x=0$ 两侧函数都是凹向下的. 另一方面, $x=7$ 却是拐点, 因为在 7 的左边函数是凹向下的, 而在 7 的右边是凹向上的.

正如我们在上一节提醒的, 第三行中的图示旨在作为你作图时的导引. 它们表明原始函数在哪里是凹向上的, 在哪里是凹向下的. 对于函数图像实际上是什么样子的, 它们只能给个大概. 这正是为什么我们需要去看一种绘制函数图像的全面方法. 前面提到的三种类型的符号表格会在这种方法中用到, 但事情远不止于此. 所以系好安全带, 我们现在出发 ……

12.2　绘制函数图像的全面方法

下边是一个绘制函数图像的十一步方法. 在你开始绘制图像前, 请先画好坐标轴, 这样就能把收集到的一些关键信息标记在图像上.

(1) **对称性**　通过用 $-x$ 替换 x, 然后看是否能得到原始函数, 来检验函数是奇函数、偶函数或者两者都不是. 如果函数奇函数或偶函数, 你只需画出 $x \geqslant 0$ 的部分, 另外一部分可以通过对称性得到. 这能为你节省很多时间.

(2) **y 轴截距**　通过设 $x = 0$ 来求 y 轴截距 (如果存在的话), 并把它标记在图像上.

(3) **x 轴截距**　通过设 $y = 0$ 并解得 x 来求 x 轴截距 x. 但这有时会很困难, 甚至不可能. 例如, 如果要因式分解一个次数为三或更高的多项式, 可能需要反复观察找出一个根, 然后利用多项式除法降次, 再继续因式分解. 在图像上标记 x 轴截距.

(4) **定义域**　求出函数 f 的定义域. 如果定义域在 f 的定义中已给出, 那不需要再做什么; 否则的话, 定义域应该包括实数线上尽可能多的部分. 记住, 要剔除那些使得分母为 0、偶次根号下的量为负数, 或者对数符号里的量为负数或 0 的数. 如果牵扯到反三角函数, 情况就更复杂了, 所以我建议你记住所有反三角函数的定义域. (例如, 无法取不在区间 $[-1,1]$ 中的数的反正弦函数.)

(5) **垂直渐近线**　它们通常出现在分母为 0 的位置. (如果有分母的话!) 注意: 如果此时的分子也为零, 那得到的是一个可去不连续点[①] 而不是一条垂直渐近线. 此外, 也可能由于对数因式而得到垂直渐近线. 在图像上用垂直的虚线来标记所有的垂直渐近线.

(6) **函数的正负**　像 12.1 节描述的那样建立一个符号表格. 从上边的第 (3) 步可知函数的零点, 从第 (4) 步和第 (5) 步可知函数的不连续点. 这个表格会告诉你, 在哪里函数图像位于 x 轴之上, 在哪里位于 x 轴之下.

(7) **水平渐近线**　通过计算 $\lim\limits_{x\to\infty} f(x)$ 和 $\lim\limits_{x\to-\infty} f(x)$来找出函数的水平渐近线. 即使这个极限为 $\pm\infty$, 它也会告诉你当 x 非常大 (或负的非常大) 时函数的走势, 从而得到某种 “倾斜” 渐近线. 不管怎样, 如果有水平渐近线, 用水平的虚线在图像中标记出来. 在这里, 你可以在水平和垂直渐近线周围选取一些合适的点去计算这些点的函数值, 并制成符号表格, 以此来判断函数图像位于渐近线的哪一侧.

(8) **导数的正负**　现在轮到微积分上场了. 求出一阶导数, 找到所有的临界点. 回想一下, 临界点是导数为 0 的点或导数不存在的点. 像 12.1.1 节讲解的那样, 绘制一个关于一阶导数的符号表格. 从表格的第三行了解该函数何时为增函数, 或者

① 例如, 如果 $f(x) = (x^2 - 3x + 2)/(x - 2)$. 通过因式分解, 分子变为 $(x - 1)(x - 2)$, 很容易看出 $f(x) = x - 1$ (除去 $x = 2$, 在那里函数 f 没有定义). 其图像可见 3.1 节.

何时为减函数, 何时为水平.

(9) **最大值和最小值** 从上面的符号表格中, 你能找到所有的局部最大值或最小值. 回想一下, 这些值仅出现在临界点处. 对于每一个最大值和最小值, 你都需要把 x 的值代入 $y = f(x)$, 求出对应的函数值. 要确保你把这些点标记在了函数图像上.

(10) **二阶导数的正负** 求出二阶导数, 并找到所有二阶导数为零或不存在的点. 像 12.1.2 节描述的那样, 绘制一个关于二阶导数的符号表格. 该表格的第三行说明了函数图像在哪里是凹向上的, 又在哪里是凹向下的.

(11) **拐点** 使用二阶导数的符号表格去寻找拐点. 回想一下, 在拐点处的二阶导数一定为 0, 并在该点的两侧二阶导数的符号是相反的. 对于每一个拐点 x, 你都需要将其代入 $y = f(x)$ 来求出对应的函数值, 并把这些点标记在图像上.

现在, 使用所有你收集到的信息去完成函数图像的绘制. 如果哪里出现了不一致, 那你可能什么地方出错了! 你收集到的所有这些信息应该是能够严丝合缝地拼凑在一起的漂亮的函数图像.

顺便提一下, 对于第 (9) 步的局部最大值和最小值, 记住你也可以使用二阶导数的正负去找 (参见 11.5.2 节). 不过, 这个方法有时并不适用 —— 这也正是我推荐使用一阶导数的符号表格的原因.

12.3 例　题

我们先看一个不使用一阶导数和二阶导数的例子, 再看四个使用上述全面方法的例子.

12.3.1 一个不使用导数的例子



在 12.1 节的开始, 我们提到过函数

$$f(x) = \frac{(x-3)(x-1)^2}{x^3(x+2)}.$$

现在让我们仅用上述程序的前七步去绘制函数图像.

(1) **对称性** 把 $-x$ 而不是 x 代入原始函数, 努力变换一番, 但这是徒劳的, 所以该函数是非奇非偶的.

(2) **y 轴截距** 设 $x = 0$, 则该函数分母为零. 所以该函数在 $x = 0$ 处趋于无穷大, 没有 y 轴截距.

(3) **x 轴截距** 设 $y = 0$, 则我们必有 $x - 3 = 0$ 或 $x - 1 = 0$, 所以 x 轴截距为 1 和 3.

(4) **定义域** 很显然, 该函数的定义域为除 0 和 -2 外的所有 x.

(5) **垂直渐近线** 当 $x=0$ 或 $x=-2$ 时, 分母都趋于 0, 而此时分子不为 0, 所以这两处有垂直渐近线.

(6) **函数的正负** 这一点已经深入讨论过了, 我们知道该函数在 $(-2,0)$ 和 $(3,\infty)$ 为正, 其余全为负 (除了在 x 轴截距和垂直渐近线处). 作为参考, 图 12-11 是在 12.1 节中出现的表格.

x	-3	-2	-1	0	$\frac{1}{2}$	1	2	3	4
$f(x)$	$-$	$\star$	$+$	$\star$	$-$	0	$-$	0	$+$

图 12-11

(7) **水平渐近线** 为此, 需要去求

$$\lim_{x\to\infty}\frac{(x-3)(x-1)^2}{x^3(x+2)} \text{ 和 } \lim_{x\to-\infty}\frac{(x-3)(x-1)^2}{x^3(x+2)}.$$

我留给你来证明这两个极限均为 0 (使用 4.3 节中的方法), 所以该函数在 $y=0$ 有一条双侧水平渐近线.

现在可以画函数图像了. 让我们先把已知的点标记在图 12-12 上.

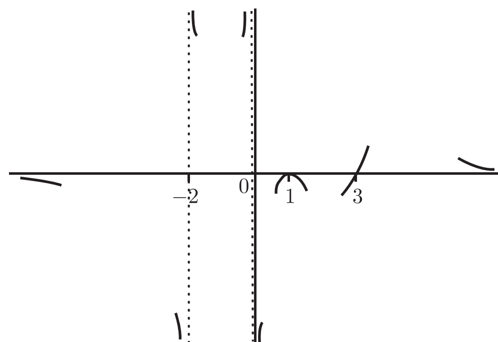


图 12-12

两条水平渐近线都是 $y=0$. 在图像的左手边, 曲线在 x 轴的下方, 因为当 $x<-2$ 时函数值是负的. 在图像的右手边, 曲线在 x 轴的上方, 因为当 $x>3$ 时函数值是正的 (可通过符号表格看出来). 至于垂直渐近线, 在 $x=-2$ 的垂直渐近线的右侧, 函数为正, 在其左侧则为负 (再次用到了符号表格). 用同样的方式来分析 $x=0$ 的垂直渐近线. 现在考虑 x 轴截距. 在 $x=1$ 点函数与 x 轴相切, 因为在该点的两侧函数值都是负的. 另一方面, 在另一点 $x=3$, 函数通过 x 轴, 因为在该点两侧的函数值的正负是相反的. 下面让我们把这些小段用平滑的曲线连接起来, 从而得到图 12-13.

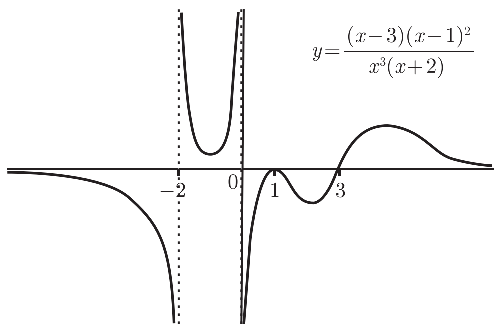


图 12-13

这是对函数图像大致样子的一个相当不错的近似. 可问题是, 除了知道在 $x = 1$ 有局部最大值, 我们并不知道其他的局部最大值和最小值. 显然, 在 $x = -2$ 和 $x = 0$ 之间有至少一个局部最小值, 在 $x = 1$ 和 $x = 3$ 之间有至少一个局部最小值, 在 $x = 3$ 右边有至少一个局部最大值. 不过, 极值还可能有更多 —— 图像可能有比目前展现的更多的起伏. 不使用导数我们是无法判断的.

那么为什么不使用导数呢? 因为对于这个函数, 导数实在太难求解了! 如果你不怕麻烦去求导, 就会发现

$$f'(x) = \frac{-x^4 + 10x^3 - 11x^2 - 16x + 18}{x^4(x+2)^2}.$$

我们实际上已经知道 $x = 1$ 是它的局部最大值, 所以 $f'(1)$ 应该为 0. 你可以检验一下, 发现当 x =1 时, 分子确实为 0. 这意味着 $(x-1)$ 为分子的一个因式, 通过长除法, 可以发现分子为 $(x-1)(-x^3 + 9x^2 - 2x - 18)$. 这仍然留下了一个三次方程需要去处理, 但至少我们知道这个三次方程最多有三个解. 这意味着除了 $x = 1$ 外, 最多还有另外三个临界点. 具体说, 这个图像并没有更多的起伏, 有的只是从上图中可以看出的四个临界点.

至于使用二阶导数去找出凹性和拐点, 我只能说, 情况会比一阶导数的还要糟糕. 所幸另一方面, 并不是每一个函数都有这么难以处理的导数 —— 让我们看以下的四个例子, 在它们身上可以应用完整的方法.

12.3.2 完整的方法：例一

在 11.5.1 节的结尾部分, 我们看到函数 $f(x) = x\ln(x)$ 在 $x = 1/\mathrm{e}$ 这点有局部最小值. 我们甚至画出了它的局部图像. 现在就应用完整的方法把函数 $y = f(x)$ 的图像补充完整.



(1) **对称性** 当 $x \leqslant 0$ 时, 该函数甚至没有定义, 所以它显然不可能是奇函数或偶函数.

(2) ***y* 轴截距** 设 $x = 0$, 则该函数在 x =0 没有定义, 所以它不可能有 y 轴

截距.

(3) **x 轴截距** 设 $y=0$, 则我们必有 $x=0$ 或 $\ln(x)=0$. 不可能有 $x=0$, 因为在 $x=0$ 处没有定义; 如果 $\ln(x)=0$, 那么 $x=1$. 所以唯一的 x 轴截距为 $x=1$.

(4) **定义域** 由于有因子 $\ln(x)$, 所以该函数的定义域必定为 $(0, +\infty)$.

(5) **垂直渐近线** 因子 $\ln(x)$ 是否可能会在 $x=0$ 引入一条垂直渐近线? 让我们检验一下. 由于该函数只有在 $x>0$ 才有定义, 所以只需要考虑它的右极限, 即 $\lim\limits_{x\to 0^+} x\ln(x)$. 实际上, 从 9.4.6 节我们已知, 这个极限为 0, 因为随着 $x\to 0^+$ 对数函数缓慢地趋于 $-\infty$. 所以该函数没有垂直渐近线, 仅仅在原点有个 (右侧) 可去不连续点.

(6) **函数的正负** 我们已经知道该函数对于 $x\leqslant 0$ 没有定义, 与 x 轴的截距仅仅有一点 $x=1$. 所以还需要在其之间的空格填入诸如 $x=1/2$ 和 $x=2$. 当 $x=1/2$ 时, $\ln(1/2)=-\ln(2)$, 为负, 所以 f 的符号为 $(-)$. 当 $x=2$ 时, 很容易可以看出 f 的符号为 $(+)$. 这样符号表格看上去如图 12-14.

x	$\leqslant 0$	$\frac{1}{2}$	1	2
$f(x)$	$\star$	$-$	0	$+$

图 12-14

(7) **水平渐近线** 仅需要考虑 $\lim\limits_{x\to\infty} x\ln(x)$, 因为 $x\to-\infty$ 的极限甚至说都说不通. 而上述极限显然为 ∞, 因为随着 $x\to\infty$, x 和 $\ln(x)$ 都趋于 ∞. 所以也没有水平渐近线.

(8) **导数的正负** 通过使用乘积法则, 可以得出 $f'(x)=\ln(x)+1$ (正如在 11.5.1 节计算过的). 所以当 $\ln(x)=-1$, 即 $x=\mathrm{e}^{-1}=1/\mathrm{e}$ 时, $f'(x)=0$. 我们只需选在 $x=0$ 和 $x=1/\mathrm{e}$ 之间的一点, 以及大于 $x=1/\mathrm{e}$ 的另一点. 不妨分别选 $x=1/10$ 和 $x=1$. 注意到 $f'(1/10)=\ln(1/10)+1=-\ln(10)+1$, 它显然为负; 而 $f'(1)=\ln(1)+1$, 它为正. 这样, $f'(x)$ 的符号表格看上去如图 12-15.

x	$\leqslant 0$	$\frac{1}{10}$	$\frac{1}{\mathrm{e}}$	1
$f'(x)$	$\star$	$-$	0	$+$
		\	—	/

图 12-15

(9) **最大值和最小值** 通过上边的表格可以知道, 仅仅在 $x=1/\mathrm{e}$ 点有局部最小值. 现在只需计算出 y 值: $y=\mathrm{e}^{-1}\ln(\mathrm{e}^{-1})=-\mathrm{e}^{-1}=-1/\mathrm{e}$. 所以局部最小值的坐标为 $(1/\mathrm{e}, -1/\mathrm{e})$, 正如我们在 11.5.1 节已经见到的那样.

(10) **二阶导数的正负** 由于 $f'(x)=\ln(x)+1$, 有 $f''(x)=1/x$. 又由于函数 f 的定义域为 $x>0$, 所以对于相关的 x, 都有 $f''(x)>0$. 这意味着 f 始终是凹向上的.

(11) **拐点** 由于 $f''(x)=1/x$, 永远不可能为 0, 所以没有拐点.

现在把所有收集到的信息标记在图像上. 我们在原点有一个可去不连续点, 在点 $(1/\mathrm{e}, -1/\mathrm{e})$ 有一个局部最小值, 在 x 轴上的截距为 1, 并且没有水平或垂直渐近线.

当 x <1 时, 图像在 x 轴的下方; 当 x >1 时, 图像在 x 轴的上方. 此外, 函数当 $0 < x < 1/\mathrm{e}$ 时为减函数, 当 $x > 1/\mathrm{e}$ 时为增函数, 并且始终是凹向上的. 因此, 它的图像必定看上去如图 12-16.

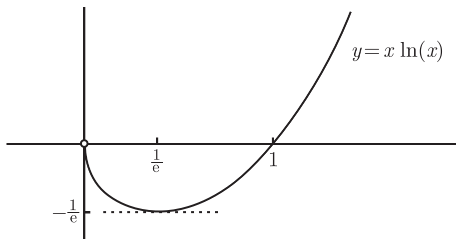


图 12-16

这不能说完美, 但比起 11.5.1 节的初步尝试是要好太多了, 毕竟我们现在知道了多得多的信息.

12.3.3 完整的方法：例二

再看一个以前的例子：$f(x) = x^2(x-5)^3$. 在 10.1.4 节中, 已经绘制出了 $y = f(x)$ 图像的大致样子; 而在 12.1 节中, 也已经制作了 $f(x)$、$f'(x)$ 和 $f''(x)$ 的符号表格. 这意味着我们可以加大油门, 快速通过.



(1) **对称性** 如果用 $(-x)$ 替换 x, 你会得到 $f(-x) = (-x)^2(-x-5)^3 = -x^2(x+5)^3$. 这既不是 $f(x)$ 也不是 $-f(x)$, 所以该函数非奇非偶. 好吧, 你总不可能事事顺心如意.

(2) **y 轴截距** 当 x =0 时, $y = f(0) = 0$. 所以 y 轴截距为 $y = 0$.

(3) **x 轴截距** 如果 y =0, 则肯定有 $x^2 = 0$ 或 $(x-5)^3 = 0$, 所以 x 轴截距为 $x = 0$ 或 $x = 5$.

(4) **定义域** 很显然, $f(x)$ 可以取任意的 x, 所以该函数的定义域为全体实数 $\mathbb{R}$.

(5) **垂直渐近线** 由于定义域为全体实数, 所以没有垂直渐近线.

(6) **函数的正负** 正如我们在 12.1 节所见, $f(x)$ 的符号表格如图 12-17.

x	-1	0	2	5	6
$f(x)$	$-$	0	$-$	0	$+$

图 12-17

所以仅当 $x > 5$ 时, 图像在 x 轴上方.

(7) **水平渐近线** 很容易看出：

$$\lim_{x\to\infty} x^2\left(x-5\right)^3 = \infty \text{ 和 } \lim_{x\to-\infty} x^2\left(x-5\right)^3 = -\infty.$$

毕竟, 当 $x \to \infty$ 时, x^2 和 $(x-5)^3$ 都趋于 ∞, 因此它们的乘积也趋于 ∞; 而当 $x \to -\infty$ 时, x^2 趋于 ∞ 而 $(x-5)^3$ 趋于 $-\infty$, 所以乘积趋于 $-\infty$. 我们还可能注意到, 当 x 非常大 (正的或负的) 时, 量 $(x-5)$ 表现得像其最高次项 x, 所以在图像的左右两端 (但不是在原点附近), $x^2(x-5)^3$ 表现得像 x^5.

(8) **导数的正负**　正如我们在 12.1.1 节所见, $f'(x)$ 的符号表格如图 12-18.

x	-1	0	1	2	3	5	6
$f'(x)$	$+$	0	$-$	0	$+$	0	$+$
	/	—	\	_	/	—	/

图　12-18

这告诉了我们函数在哪里为增函数、在哪里为减函数, 或在哪里为水平的.

(9) **最大值和最小值**　从上表中可以看出: $x=0$ 为局部最大值, $x=2$ 是局部最小值, 而 $x=5$ 是水平拐点. 现在需要计算这些点对应的函数值. 通过把这些 x 值代入 $f(x)=x^2(x-5)^3$ 可得: $f(0)=0$, $f(2)=(2)^2(-3)^3=-108$, 以及 $f(5)=0$. 所以在原点处有局部最大值, 在点 $(2,-108)$ 有局部最小值, 而点 $(5, 0)$ 是水平拐点.

(10) **二阶导数的正负**　在 12.1.2 节中, 我们已经见到图 12-19.

x	0	$2-\frac{1}{2}\sqrt{6}$	2	$2+\frac{1}{2}\sqrt{6}$	4	5	6
$f''(x)$	$-$	0	$+$	0	$-$	0	$+$
	⌢	·	⌣	·	⌢	·	⌣

图　12-19

通过该表格, 可以看出函数在哪里是凹向上, 在哪里是凹向下. 注意到 $f''(0)<0$ 以及 $f''(2)>0$, 前者再次确认了临界点 $x=0$ 是局部最大值, 而后者则再次确认了临界点 $x=2$ 是局部最小值.

(11) **拐点**　从上表中还可以判断出 $x=2-\dfrac{1}{2}\sqrt{6}, x=2+\dfrac{1}{2}\sqrt{6}$ 和 $x=5$ 为该函数的拐点. 事实上, 最后一个点是我们早就知道的, 因为在步骤 (9) 已经看到点 (5, 0) 为水平拐点. 其他两个点则要麻烦许多, 需要分别把 $x=2-\dfrac{1}{2}\sqrt{6}$, $x=2+\dfrac{1}{2}\sqrt{6}$ 代入原始函数 $f(x)=x^2(x-5)^3$. 不幸的是, 得到的结果一团糟. 这里我们取个巧, 设 $\alpha=f\left(2-\dfrac{1}{2}\sqrt{6}\right)$ 和 $\beta=f\left(2+\dfrac{1}{2}\sqrt{6}\right)$. 这意味着

$$\alpha=\left(2-\frac{1}{2}\sqrt{6}\right)^2\left(-3-\frac{1}{2}\sqrt{6}\right)^3, \beta=\left(2+\frac{1}{2}\sqrt{6}\right)^2\left(-3+\frac{1}{2}\sqrt{6}\right)^3.$$

实际上, 如果费劲把它乘开, 你可以化简这个表达式, 但这毫无乐趣可言. 我们也可以难得使用计算器去计算得到 α 大约等于 -45.3, β 大约等于 -58.2. 但这些仅仅

是近似值! 计算器不可能给出像 α或 β 这样的无理数的准确数值. 不管怎样, 我们知道该函数的拐点为 $\left(2-\dfrac{1}{2}\sqrt{6},\,\alpha\right)$, $\left(2+\dfrac{1}{2}\sqrt{6},\,\beta\right)$ 和 $(5,\,0)$.

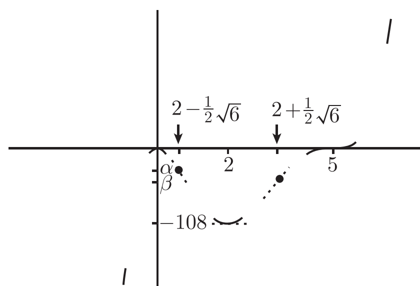


图 12-20

现在, 让我们把一切拼凑起来. 画出坐标系, 标注原点处的 y 轴截距、0 和 5 处的 x 轴截距, 原点处的局部最大值、$(2,\,-108)$ 处的局部最小值, 以及 $(5,\,0)$ 处的水平拐点、$\left(2-\dfrac{1}{2}\sqrt{6},\,\alpha\right)$和 $\left(2+\dfrac{1}{2}\sqrt{6},\,\beta\right)$ 处的非水平拐点. 我们还知道当 $x\to\infty$ 时, $y\to\infty$ 以及 $x\to-\infty$ 时, $y\to-\infty$, 所以可以用小段曲线表示这一点. 综合起来, 得到图 12-20.

注意到我们从 $f'(x)$ 的符号表格已知, 曲线在拐点 $\left(2-\dfrac{1}{2}\sqrt{6}\right)$ 处的斜率为负, 在拐点 $\left(2+\dfrac{1}{2}\sqrt{6}\right)$ 处的斜率为正. 现在只需把各段连接起来, 得到图 12-21.

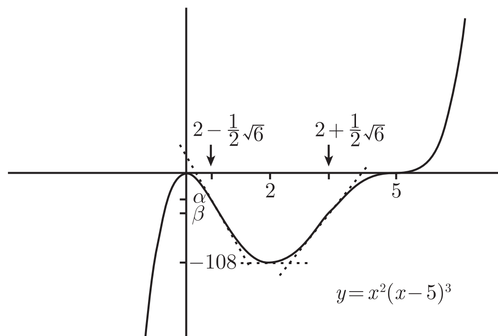


图 12-21

再一次地, 这比我们在 10.1.4 节中的绘图尝试要做得更好, 因为它还表示出了拐点.

12.3.4 完整的方法：例三

现在, 让我们绘制 $y=f(x)$ 的函数图像, 其中

$$f(x)=x\mathrm{e}^{-3x^2/2}.$$

(1) **对称性** 用 $(-x)$ 代替 x, 我们得到 $-x\mathrm{e}^{-3(-x)^2/2}=-x\mathrm{e}^{-3x^2/2}=-f(x)$, 所以该函数为奇函数. 这是一个意外之喜：仅仅需要绘制 $x\geqslant 0$ 的部分, 剩下的一半



则很容易得到.

(2) **y 轴截距**　当 $x=$0 时, $y=0\mathrm{e}^{-3(0)^2/2}=0$. 所以 y 轴截距为 $y=0$.

(3) **x 轴截距**　当 $y=0$ 时, $0=x\mathrm{e}^{-3x^2/2}$, 所以要么 $x=0$ 要么 $\mathrm{e}^{-3x^2/2}=0$. 后一个方程是无解的, 因为指数函数永远为正. 因此 x 轴截距仅有 $x=0$. 到目前为止, 我们知道的只是, 该函数为奇函数且它与坐标轴只相交于原点.

(4) **定义域**　很明显, x 可以取任意值而不会引出问题 —— 这里没有偶次根或对数, 而即使把函数写成

$$y=\frac{x}{\mathrm{e}^{3x^2/2}},$$

分母也不会为 0, 因为指数函数始终为正. 所以定义域为全体实数 $\mathbb{R}$.

(5) **垂直渐近线**　没有垂直渐近线, 因为定义域为全体实数 $\mathbb{R}$.

(6) **函数的正负**　我们知道使 $f(x)=0$ 的点仅有一点, 就是 $x=0$ 时. 这样就有图 12-22 所示的这个极其简单的符号表格.

x	-1	0	1
$f(x)$	$-$	0	$+$

图　12-22

从表格中可以看出, 当 $x>0$ 时函数为正, 当 $x<0$ 时函数为负.

(7) **水平渐近线**　为此, 需要求

$$\lim_{x\to\infty}\frac{x}{\mathrm{e}^{3x^2/2}}\text{和}\lim_{x\to-\infty}\frac{x}{\mathrm{e}^{3x^2/2}}$$

注意到在两种情况下 $3x^2/2$ 都是一个很大的正数, 所以分母是一个很大的指数. 由于指数函数增长迅速 (参见 9.4.4 节), 两个极限都为 0. 所以有一条双侧水平渐近线 $y=0$.

(8) **导数的正负**　现在需要求导. 通过使用乘积法则和链式求导法则, 你可以检验一下

$$f'(x)=x(-3x)\mathrm{e}^{-3x^2/2}+\mathrm{e}^{-3x^2/2}=(1-3x^2)\mathrm{e}^{-3x^2/2}.$$

它处处都有定义, 但它什么时候为 0 呢? 由于指数函数始终为正, 所以仅当 $1-3x^2=0$, 即当 $x=1/\sqrt{3}$ 或 $x=-1/\sqrt{3}$ 时, 导数才为 0. 让我们选 -1, 0 和 1 填充之间的空白, 这样导数的符号表格看上去像图 12-23.

x	-1	$\frac{-1}{\sqrt{3}}$	0	$\frac{1}{\sqrt{3}}$	1
$f'(x)$	$-$	0	$+$	0	$-$
	\	_	/	‾	\

图　12-23

从表格中可以看出, 函数在 $-1/\sqrt{3}$ 和 $\sqrt{3}$ 之间时为增函数, 在其他区间则为减函数. 还注意到 f 是奇函数这一点 (从步骤 (1) 可知) 在上表第三行中显而易见.

(9) **最大值和最小值** 从上面的符号表格很容易看出, $x = 1/\sqrt{3}$ 对应的点是局部最大值, $x = -1/\sqrt{3}$ 对应的点为局部最小值. 剩下要做的只是把 x 值分别代入原始函数以求得 y 值. 当 $x = 1/\sqrt{3}$ 时, 有

$$y = \frac{1}{\sqrt{3}}\mathrm{e}^{-3(1/\sqrt{3})^2/2} = \frac{\mathrm{e}^{-1/2}}{\sqrt{3}}.$$

这样在点 $\left(1/\sqrt{3}, \mathrm{e}^{-1/2}/\sqrt{3}\right)$ 有局部最大值. 又由于函数为奇函数, 我们甚至不需要把 $x = -1/\sqrt{3}$ 代入就可以看出, 必有 $(-1/\sqrt{3}, -\mathrm{e}^{-1/2}/\sqrt{3})$ 为局部最小值.

(10) **二阶导数的正负** 为此需要再次求导, 再次使用乘积法则和链式求导法则. 得到

$$f''(x) = (1 - 3x^2)(-3x)\mathrm{e}^{-3x^2/2} + (-6x)\mathrm{e}^{-3x^2/2} = 9x(x^2 - 1)\mathrm{e}^{-3x^2/2}.$$

再一次地, 由于指数函数始终为正, 所以仅当 $x = 0$ 或 $x^2 - 1 = 0$, 即当 $x =0$, $x =1$ 或 $x = -1$ 时, $f''(x)$ 才为零. 其符号表格看上去如图 12-24.

x	-2	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1	2
$f''(x)$	$-$	0	$+$	0	$-$	0	$+$
	⌢	·	⌣	·	⌢	·	⌣

图　12-24

当 $x =1/2$ 时, 因子 $9x$ 为正, 但 $(x^2 - 1)$ 为负, 同时指数函数始终为正, 所以整个结果为负. 当 $x =2$ 时, 同样容易看出二阶导数为正. $x = -1/2$ 和 $x = -2$ 时的情况也容易判断, 并且遵从对称性. (由于原始函数为奇函数, 它的导函数为偶函数, 它的二阶导函数为奇函数. 你可能需要在这上面稍微思考一下!) 从第三行可以看出, 当 $x < -1$ 或 $0 < x < 1$ 时, 函数的图像是凹向下的; 当 $x >1$ 或 $-1 < x <0$ 时, 图像是凹向上的. 顺便说一下, 注意到在临界点 $x = 1/\sqrt{3}$, 二阶导数为负 —— 这再次确认了在该点有局部最大值. 类似地, 当 $x = -1/\sqrt{3}$, 二阶导数为正, 所以确实在该点有局部最小值.

(11) **拐点** 从上表中可以看出, 在 $x =0$, $x =1$ 或 $x = -1$ 这些点, 函数的凹性都会发生变化. 所以这些点都是函数的拐点, 而我们需要做的仅是求出这些点对应的 y 值. 通过把这些点代入原始函数 $y = x\mathrm{e}^{-3x^2/2}$, 容易得到这些拐点的坐标分别为 $\left(1, \mathrm{e}^{-3/2}\right), \left(-1, -\mathrm{e}^{-3/2}\right)$ 和 $(0, 0)$.

如果你真的一直很乖的话, 应该已经在坐标系上标出了所有已知的信息, 得到图 12-25.

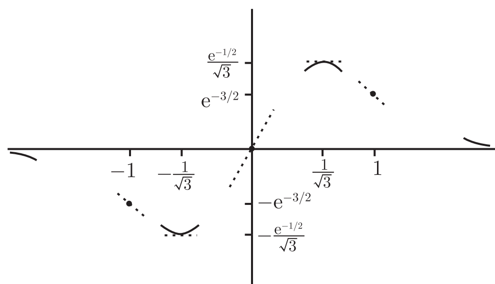


图　12-25

在上图中可以看到 x 轴和 y 轴截距 (都在原点)、水平渐近线 (x 轴)、在 $\left(1/\sqrt{3}, e^{-1/2}/\sqrt{3}\right)$ 的最大值、在 $\left(-1/\sqrt{3}, -e^{-1/2}/\sqrt{3}\right)$ 的最小值, 以及在 (0, 0) 和 $\left(1, e^{-3/2}\right), \left(-1, -e^{-3/2}\right)$ 的拐点 (在图中暂时用虚线表示). 由于在步骤 (6) 知道了 $f(x)$ 的正负, 甚至已经分析了该函数在水平渐近线附近的走势, 所以在图中将这一点体现了出来. 不管怎样, 剩下需要做的只是如图 12-26 把各段连接起来. 这样就确实把图像的所有重要特征都体现了出来.

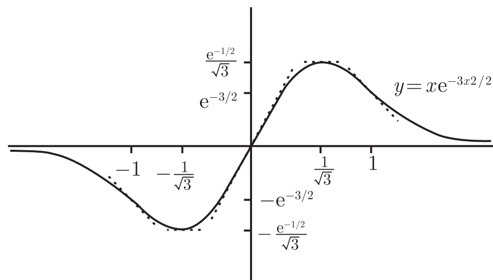


图　12-26

12.3.5　完整的方法：例四



再举一个例子：画 $y = f(x)$ 的图像, 其中 f 是令人望而生畏的

$$f(x) = \frac{x^3 - 6x^2 + 13x - 8}{x}.$$

(1) **对称性**　用 $(-x)$ 代替 x, 我们得到 $\left(-x^3 - 6x^2 - 13x - 8\right)/(-x)$, 它既不是 $f(x)$ 也不是 $-f(x)$, 所以该函数没有对称性. 真遗憾.

(2) **y 轴截距**　把 x =0 代入, 我们得到 $-8/0$, 这是没有定义的. 所以没有 y 轴截距.

(3) **x 轴截距**　情况变得麻烦了. 需要设 $y = 0$, 这意味着 $x^3 - 6x^2 + 13x - 8 = 0$. 这是一个三次方程, 所以因式分解可能会让人头疼. 最好的办法是试根. 试试 $x = 1$. 得到 $1 - 6 + 13 - 8 = 0$, 猜对了! (基本上说, 简单的根应该是常数项 -8 的因子, 所以如果 $\pm 1, \pm 2, \pm 4$ 和 ± 8 都不适用, 那你就完蛋了.)　所幸, 我们的初次尝试就成

功了, 我们知道 $(x-1)$ 为它的一个因子. 接下来, 做多项式除法：

$$x-1\overline{)x^3-6x^2+13x-8}$$



我留给你来完成这个除法, 并得到另一个因子为 x^2-5x+8. 你再能对这个二次函数因式分解吗? 它的判别式为 $(-5)^2-4(8)=-7$, 为负, 所以你不能进行因式分解. 也就是说, 有 $x^3-6x^2+13x-8=(x-1)\left(x^2-5x+8\right)$. 由于第二个因子始终为正, 所以在 x 轴仅有的截距为 $x=1$.

(4) **定义域**　仅有的问题是 $x=0$, 所以定义域为 $\mathbb{R}\backslash\{0\}$.

(5) **垂直渐近线**　$x=0$ 处有垂直渐近线, 因为此时分母为 0, 而分子不为 0. 不可能再有其他的垂直渐近线, 因为函数在其他地方处处有定义.

(6) **函数的正负**　把函数写为

$$f(x)=\frac{(x-1)(x^2-5x+8)}{x}.$$

在 x 轴的唯一截距为 $x=1$, 唯一的不连续点在 $x=0$ 处, 所以符号表格看上去像图 12-27. (确认你理解了在 $x=-1$, $x=1/2$ 和 $x=2$ 时函数的正负.)

x	-1	0	$\frac{1}{2}$	1	2
$f(x)$	$+$	$\star$	$-$	0	$+$

图　12-27

(7) **水平渐近线**　考虑极限

$$\lim_{x\to\infty}\frac{x^3-6x^2+13x-8}{x}\quad 和\quad \lim_{x\to-\infty}\frac{x^3-6x^2+13x-8}{x}.$$

它们可以改写为

$$\lim_{x\to\infty}\left(x^2-6x+13-\frac{8}{x}\right)\quad 和\quad \lim_{x\to-\infty}\left(x^2-6x+13-\frac{8}{x}\right).$$

很明显, 这两个极限的结果都是无穷大, 所以没有水平渐近线. 另一方面, 当 x 非常大 (正的或负的) 时, $f(x)$ 表现得像其主导项, 即 x^2. 所以当 x 非常大时, 曲线应该看上去很像抛物线 $y=x^2$, 如图 12-28 所示. 不管怎样, 尽管我们还没有求导, 仍旧知道了这个函数的很多信息.

注意到我们刚才用了 $f(x)$ 的符号表格去判断图像在垂直渐近线附近的情况. 具体说, 当 x 比 0 稍小时, 函数值为正, 所以曲线在垂直渐近线的左边趋于 ∞. 类似地, 当 x 比 0 略大时, 函数值为负, 这意味着曲线在垂直渐近线的右边趋于 $-\infty$.

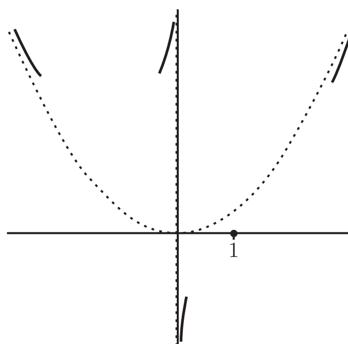


图　12-28

(8) **导数的正负**　我们已经用到了 $f(x)$ 的三种形式：

$$f(x)=\frac{x^3-6x^2+13x-8}{x}=\frac{(x-1)(x^2-5x+8)}{x}=x^2-6x+13-\frac{8}{x}.$$

为了求 $f'(x)$, 可以选 $f(x)$ 的任意一种形式. 我选第三种形式, 因为它的求导无须使用乘积法则或商法则. 于是有

$$f'(x)=2x-6+\frac{8}{x^2},$$

它可改写为

$$f'(x)=\frac{2x^3-6x^2+8}{x^2}.$$

那么何时导数为零? 何时导数不存在呢? 很显然, 当 $x=0$ 时, 导数不存在. 另一方面, 如果 $f'(x)=0$, 我们肯定有 $2x^3-6x^2+8=0$. 再一次地, 我们需要猜出三次方程的一个根. 这一次, $x=1$ 不适用, 所以再试一下 $x=-1$. 啊哈, 刚刚好! 在做完长除法后, 就可以将三次方程因式分解为 $2(x+1)(x-2)^2$. 也就是说,

$$f'(x)=\frac{2(x+1)(x-2)^2}{x^2}.$$

所以导数在 $x=0$ 处没有定义, 当 $x=-1$ 或 $x=2$ 时为 0. 现在, 可以画出 $f'(x)$ 的符号表格如图 12-29.

x	-2	-1	$-\frac{1}{2}$	0	1	2	3
$f'(x)$	$-$	0	$+$	$\star$	$+$	0	$+$
	\	_	/	⋮	/	—	/

图　12-29

确认你检验了该表格的各个细节. 从表中可以看出：当 $x>-1$ 时, 函数为增函数 (除了 $x=0$ 和 $x=2$ 这两个临界点); 当 $x<-1$ 时, 函数为减函数.

(9) **最大值和最小值**　从上表可以看出, $x=-1$ 为局部最小值, $x=2$ 为水平拐点. 现在要求对应的 y 值; 不难看出 $f(-1)=28$ 和 $f(2)=1$. 所以点 $(-1,28)$ 为

局部最小值, 点 (2, 1) 为水平拐点.

(10) **二阶导数的正负**　我们已知 $x=2$ 为一个拐点, 还有其他拐点吗? 找找看. 从形式

$$f'(x)=2x-6+\frac{8}{x^2}$$

得出

$$f''(x)=2-\frac{16}{x^3}=\frac{2(x^3-8)}{x^3}.$$

二阶导数在 $x=0$ 处没有定义, 当 $x^3-8=0$, 即 $x=2$ 时为 0. 所以没有其他拐点了! 让我们绘制其符号表格如图 12-30.

x	-1	0	1	2	3
$f''(x)$	+	$\star$	−	0	+
	$\smile$	$\vdots$	$\frown$	.	$\smile$

图　12-30

可以看到, 当 $x>2$ 和 $x<0$ 时, 图像是凹向上的; 当 $0<x<2$ 时, 图像是凹向下的. 顺便提一下, 在临界点 $x=-1$, $f''(x)>0$, 所以该点的确为局部最小值. 另一方面, 在临界点 $x=2$, $f''(2)=0$, 所以单凭它无法确认这是个拐点. 最好的确认方法是表明导数在 $x=2$ 的两侧符号相反. 而这一信息在其符号表格中清晰可见.

(11) **拐点**　我们知道 $x=2$ 是唯一的拐点, 并且知道其坐标为 (2, 1).

现在让我们基于最后几步中得到的新信息完成该函数图像的绘制. 我们需要在图像上标记出 $(-1,28)$ 处的最小值以及 (2, 1) 处的水平拐点. 但 28 是个很大的数, 所以需要压扁 y 轴 (相较于前面的草图), 以便图像的比例合适. 最后得到图 12-31.

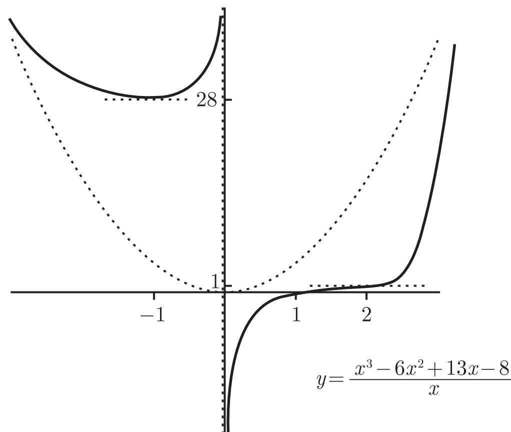


图　12-31

虚线是假想的 $y=x^2$, 虽然它的比例有点不对. 同样, 在图像的右端, 实线本该

是接近于 $y = x^2$ 的, 但我没有追求落实这一点. 不幸的是, 如果你表现了这类细节, 那不免会让拐点处的细节看不清. 确实, 图形计算器给出的结果可能像图 12-32.

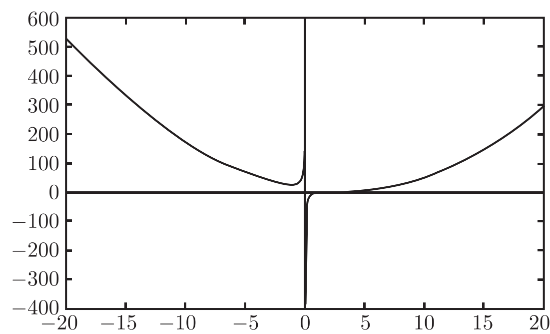


图　12-32

可以看出来曲线很像 $y = x^2$, 并且在 $x = 0$ 附近有些奇怪的表现, 但你实在看不出那里的细节. 这很好地说明了 “描点作图” 与 “作示意图” 之间的区别. 毕竟, 图形计算器只是描出足够多的点使得曲线看上去光滑, 但它没有突出图像那些有趣的特征. 如果放大图像, 你可能会对局部的细节有更好了解, 但这样你又无法顾及 x 很大时的行为了. 所以尽管前面粗略的示意图不准确, 但它对我们理解实际上发生了什么要有帮助得多, 特别是涉及极值点和拐点时：它切实展现了这些特征点都在哪里.

第 13 章　最优化和线性化

现在我们要看一下微积分的两个实际应用：最优化和线性化. 不管你相信与否, 这两个技术每天都在被工程师、经济学家及医生等用到. 简单来说, 最优化涉及找出各种可能情况中最好的一种, 不论这是在保障桥梁不会倒塌的前提下建造桥梁的最省钱方法, 还是日常如找出抵达某个目的地的最快行驶路径. 另一方面, 线性化是一种对难以计算的量找出其估算值的有用技术. 它也可被用来找出函数的零点的估算值, 这时它也被称为牛顿法. 总而言之, 我们将要讨论以下话题：

- 如何解决最优化问题, 并看三个例子;
- 使用线性化和微分估算特定的量;
- 我们的估算有多好;
- 估算函数的零点的牛顿法.

13.1　最　优　化

使某样东西 “最优化” 意味着要使之尽可能地好. 由于我们是在讨论数学, 因而这里将关注量而非质. 假设我们关心某个特定的量, 它可能是数、长度、角度、面积、成本、收入, 等等. 如果它是好的事情, 就像收入, 那么我们希望使之越大越好; 而如果它是不好的事情, 就像成本, 我们则希望使之越小越好. 简而言之, 我们想要让这个量最大化或最小化. 所以在我们的语境中, “最优化” 仅仅意味着 “相应地最大化或最小化”.

13.1.1　一个简单的最优化例子

在最近几章中, 我们已经花了相当多时间学习如何求函数的最大值和最小值. 而涉及最优化时, 通常我们关心的是全局最大值和最小值. 在 11.1.3 节, 已经提到了一个解决这个问题的很好方法. 我强烈建议你现在返回头看一下, 刷新一下记忆.



不论在哪种情况下, 都需要把这个量表示成另一个我们能控制的量的函数. 例如, 假设有两个实数的和为 10, 并且每个数都不大于 8. 那么这两个数的乘积最大可能是多少? 最小又可能是多小?

在搬出我们的方法之前, 先试探一下情况. 如果其中一个数为 8(这是它最大能取的值), 另一个数则为 2, 这时的乘积为 16. 而在另一个极端, 如果两个数都为 5, 那么乘积为 25, 显然比 16 要大. 我们能使乘积比 25 还大或比 16 还小吗? 要是这两个数分别为 $4\frac{1}{2}$ 和 $5\frac{1}{2}$ 又会怎样呢? 试试算一下.

现在, 让我们开始认真对待并设置一些变量. 假设这两个数分别为 x 和 y, 它们的乘积为 P. 于是可知有 $P=xy$. 我们想要最优化的量是 P, 但它是两个变量 x 和 y 的乘积. 这并不是我们想要的, 我们真正想要的是, 把 P 表示为一个变量的函数, 至于是其中哪个变量倒无所谓. 幸运的是, 我们还有另一个已知信息：$x+y=10$. 这意味着可以通过 $y=10-x$ 把 y 消掉. 这样的话, 就有 $P=x(10-x)$, 即把 P 表示成了单独 x 的函数.

不过, 这里有一点需要注意：P 的定义域是什么？当然, 可以在 $P=x(10-x)$ 中任意代入一个 x 值, 并得到一个有意义的答案, 但对于 x, 我们其实还知道更多 (只不过尚没有用数学语言把它描述出来)：x 不可能比 8 大. 事实上, 它也不可能比 2 小, 否则 y 就会比 8 大. 所以 x 必定位于区间 $[2, 8]$ 中. 我们应该把这视为 P 的定义域.

这样就把该文字问题重新表述为：求函数 $P=x(10-x)$ 在区间 $[2, 8]$ 上的最大值. 真不错! 我们只写出 $P=10x-x^2$, 求导得到 $\mathrm{d}P/\mathrm{d}x=10-2x$. 当 $x=5$ 时, 导数为 0, 这是唯一的临界点. 也有可能在两个端点 $x=2$ 和 $x=8$ 取到最大值或最小值. 所以潜在极值点列表包括 2, 5 和 8. 当 $x=2$ 或 $x=8$ 时, 有 $P=16$; 当 $x=5$ 时, 有 $P=25$. 因此结论是, 乘积的最大值确实为 25, 出现在两个数都为 5 的时候; 乘积的最小值确实为 16, 出现在一个数为 8 而另一个数为 2 的时候. 注意到我在陈述总结时, 并没有提到 P, x 或 y, 因为这些变量是我引入的. 如果问题中并没有给出什么变量, 那么你不仅需要识别出它们, 给它们命名, 还需要在不提及它们的前提下写出最终结论!

不妨通过图 13-1 所示的 $P'(x)$ 的符号表格[①]再检验一下刚才的结论.

x	4	5	6
$P'(x)$	$+$	0	$-$
	/	—	\

图 13-1

确实, 它是最大值. 我们也可以通过查看二阶导数的正负来验证这一点, 就像在 11.5.2 节描述的那样. 由于 $P'(x)=10-2x$, $P''(x)=-2$, 所以自然也有 $P''(5)=-2$. 由于这是负的, 再次得到 $x=5$ 为局部最大值 (它也是全局最大值). 不过, 这两种方法都不适用于端点 —— 它们仅适用于临界点.

13.1.2 最优化问题：一般方法

下边是解决最优化问题的一般方法.

(1) 识别出所有你可能用到的变量. 其中之一应该是你想要最大化或最小化的量 —— 要确保你知道是哪个! 让我们暂且把它设为 Q, 当然它也可能是其他字母, 比如 P, m 或 α.

① 参见 12.1.1 节.

(2) 试探一下当前情况的极端可能, 看变量能最大或最小到多少. (比如在上一节的例题中, x 只能在 2 和 8 之间.)

(3) 写出关联起不同变量的各个方程. 其中之一应该是关于 Q 的方程.

(4) 努力通过这些方程消去其他变量, 使得 Q 可以表示为只关于一个变量的函数.

(5) 对 Q 关于那个变量求导, 然后找出临界点. 要记住, 临界点出现在导数为 0 或不存在的位置.

(6) 求出 Q 在临界点及端点所对应的值, 从中选出最大值和最小值. 使用一阶或二阶导数的符号表格对临界点进行分类, 加以检验.

(7) 写出所得到的结论, 注意其中要用文字而非符号表示变量.

实际上, 有时候第 (4) 步可能会相当难, 不过有可能通过隐函数求导而避开这个麻烦. 我们将在 13.1.5 节看到这是如何做到的.

13.1.3 一个最优化的例子

让我们看看如何应用这个方法. 假设有个农场的边界是一道又长又直的篱笆, 而这个农场主现在想要多圈一块地来喂马. 但这个农场主有些古怪, 想圈出一个直角三角形, 并以之前的篱笆为一边 (但不是斜边), 如图 13-2 上图所示.

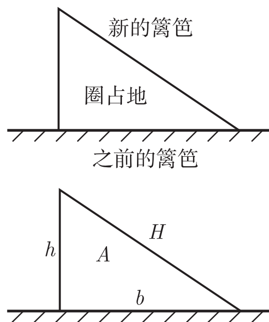


图 13-2

假设只有 300 英尺长的篱笆可供使用, 并且农场主想使新圈出的地的面积尽可能地大. 那么这块地的周长和面积分别为多少?

首先要识别出一些变量. 设三角形的底边为 b, 高为 h, 斜边为 H (单位都是英尺), 并且面积为 A (单位为平方英尺), 如图 13-2 下图所示. 注意到篱笆的长度为 $h + H$, 而我们想要最大化 A.

这样就完成了第 (1) 步. 接下来进入第 (2) 步, 考虑用 300 英尺的篱笆可以做出的极端形状, 如图 13-3 所示. 在第一种情况中, h 接近于 0, 而 b 和 H 都接近于 300, 但此时的面积很小! 在第二种情况中, b 接近于 0, 而 h 和 H 都接近于 150, 此时的面积依然非常小! 所以走中间路线, 我们应该可以做得更好. 至少已经可以确认, b 和 H 在 0 和 300 之间, 而 h 在 0 和 150 之间.

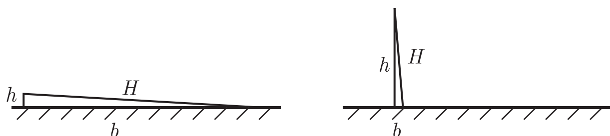


图 13-3

进入到第 (3) 步, 可以看出 $A=bh/2$ 以及 $h+H=300$. 但还需要再多一个方程, 因为需要将 b, h 和 H 这三个变量精简为一个变量. 事实上, 可以使用勾股定理, 得到 $b^2+h^2=H^2$.

现在应该试着消去一些变量. 对上式开平方, 并写出 $H=\sqrt{b^2+h^2}$ (因为 $H>0$); 将之带入 $h+H=300$, 得到 $h+\sqrt{b^2+h^2}=300$. 接下来试着把 b 消去. 从方程两边同时减去 h 再平方, 得到

$$b^2+h^2=(300-h)^2=90\,000-600h+h^2.$$

这意味着 $b=\sqrt{90\,000-600h}=10\sqrt{900-6h}$ (再一次地, 因为 b 为正, 不可能取负的平方根). 最后, 方程 $A=bh/2$ 可以被重写为

$$A=\frac{1}{2}\times 10\sqrt{900-6h}\times h=5h\sqrt{900-6h},$$

其中 h 位于区间 [0, 150]. 这样完成了第 (4) 步. 至于第 (5) 步, 可以使用乘积法则和链式求导法则, 得出

$$\frac{\mathrm{d}A}{\mathrm{d}h}=5\left(\sqrt{900-6h}+h\frac{-6}{2\sqrt{900-6h}}\right)=\frac{45(100-h)}{\sqrt{900-6h}}.$$

当 $100-h=0$, 即 $h=100$ 时, 导数为 0. 进入到第 (6) 步, 将 $h=100$ 代入上述关于 A 的方程, 得到

$$A=5(100)\sqrt{900-6(100)}=500\sqrt{300}=5000\sqrt{3}.$$

h	99	100	101
$\mathrm{d}A/\mathrm{d}h$	+	0	−
	/	—	\

图　13-4

另一方面, 对于端点 $h=0$, 我们看到 $A=0$; 类似地, 当 $h=150$ 时, 量 $900-6h$ 也趋于 0, 所以 A 又为 0. 这样得出结论：当 $h=100$ 时, A 有最大值. 我们可以用图 13-4 的导数的符号表格来检验一下. 情况还不是很坏, 因为导数的分子是 $45(100-h)$, 而分母始终为正. 所以确如我们预测的, $h=100$ 为局部最大值.

现在让我们把问题完全解决. 问题是求三角形的周长, 而现在只知道一边：$h=100$. 还要求出 b 和 H. 只需返回头去看那些方程：$h+H=300$, 所以马上有 $H=200$; 还有 $b^2+h^2=H^2$, 所以把 $h=100$ 和 $H=200$ 代入, 可知 $b=100\sqrt{3}$. 最后, 我们已经算出面积 A 的最大值为 $5000\sqrt{3}$. 所以总结陈词可以像这样：最大面积的新圈地是一个底边为 $100\sqrt{3}$ 英尺、高为 100 英尺、斜边为 200 英尺的直角三角形, 此时面积为 $5000\sqrt{3}$ 平方英尺.

13.1.4　另一个最优化的例子



下面是一个有趣的问题. 假设你要生产一批封口的、中空的圆柱体金属罐. 你可以任意选择这些罐子的尺寸, 但每个罐子的体积必须是 16π 立方英寸. 你希望使用尽可能少的金属, 因为金属的成本为 2 美分每平方英寸. 那么应该怎样选择罐子

的尺寸才能使成本最低？此时每个罐子的成本是多少？

进一步地，如果考虑到罐子的盖和底都需要焊接到罐身上，而焊接成本是 14 美分每英寸，那么情况又会如何变化？

让我们先从问题的第一部分开始. 图 13-5 是罐子的图示. 要描述圆柱体，只需要说出它的半径和高度，所以设它们分别为 r 和 h(单位为英寸). 还需要体积 V(单位为立方英寸)，因为问题也提到它了. 此外，成本取决于使用了多少金属，而这意味着圆柱体的表面积. 不妨设表面积为 A(单位为平方英寸)，成本为 C(单位为美分). 量 C 是要最小化的，虽然在这里很显然，最小化 A 等同于最小化 C. (但这对问题的第二部分就不成立了!)

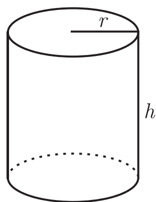


图 13-5

现在进入到第 (2) 步. 当半径 r 非常非常小时，会发生什么？这时为了满足体积必须为 16π 立方英寸的要求，高度 h 必定要非常非常大. 我们会得到图 13-6 左边那样又高又细的圆柱体. 另一方面，如果 r 非常大，那么 h 必将非常小，我们会得到图右边那样又宽又扁的圆柱体.

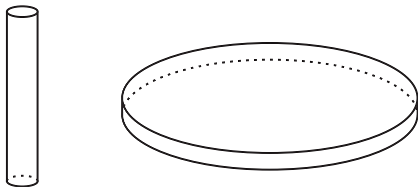


图 13-6

即使它们已经看上去相当极端，但事情可以变得更为怪异. 事实上，r 可以是任意正实数! 所以其实不存在端点，r 和 h 都位于开区间 $(0, \infty)$. 对此需要小心. 在上边的两种情况中，每个看上去都要用到很多金属，所以低成本的解决方案很可能更接近于前面那个比例协调的圆柱体，而非上面两种极端情况.

接下来进入到第 (3) 步：需要找到一些方程. 我们知道 $V = 16\pi$; 又由于圆柱体的体积 $V = \pi r^2 h$, 这样就有了第一个有用的方程：

$$16\pi = \pi r^2 h.$$

可以将它重写为 $16 = r^2 h$ 或

$$h = \frac{16}{r^2}.$$

另一方面，封口的圆柱体的表面积为

$$A = 2\pi rh + 2\pi r^2,$$

其中第一项为罐身面积, 第二项为盖和底的面积. (如果没有盖, 第二项则为 πr^2, 没有倍数 2.) 最后, 成本是 2 美分乘以总面积, 所以有

$$C = 2A = 4\pi rh + 4\pi r^2.$$

对于第 (4) 步, 注意到上述方程的右边两项都涉及 r, 所以选择消去 h 会更简单. 我们已经看到 $h = \dfrac{16}{r^2}$, 所以只需代入上述方程, 就有

$$C = 4\pi r\left(\frac{16}{r^2}\right) + 4\pi r^2 = 4\pi\left(\frac{16}{r} + r^2\right).$$

很好! 我们做到了用 r 来表示 C. 现在的问题变为, 当 r 在区间 $(0, \infty)$ 上时, 如何最小化 C. 我们有

$$\frac{\mathrm{d}C}{\mathrm{d}r} = 4\pi\left(-\frac{16}{r^2} + 2r\right),$$

它对 $(0, \infty)$ 内的所有 r 均成立, 并当 $-\dfrac{16}{r^2} + 2r = 0$ 或 $2r^3 = 16$ 时为零. 这意味着 $r^3 = 8$, 所以 $r = 2$ 是唯一的临界点. 那么端点呢? 我们无法把 $r = 0$ 代入 C 的表达式, 但可以求此时的极限:

$$\lim_{r\to 0^+} C = \lim_{r\to 0^+} 4\pi\left(\frac{16}{r} + r^2\right) = \infty.$$

该极限为无穷大, 因为当 $r \to 0^+$ 时, $\dfrac{16}{r}$ 趋于无穷大. 这意味着当半径趋于 0 时, 成本将越来越高. 这可不是我们想要的! 所以要离这个端点远远的. 那么区间 $(0, \infty)$ 的另一个端点呢? 再一次地, 我们无法令 $r = \infty$, 所以也要求极限:

$$\lim_{r\to\infty} C = \lim_{r\to\infty} 4\pi\left(\frac{16}{r} + r^2\right) = \infty.$$

这次是 r^2 项趋于无穷大. 没关系, 我们也将对这个端点敬而远之. 所以结论是, 在 $r = 2$ 点有局部最小值和全局最小值. 可以通过一阶或二阶导数的符号表格来检验一下. 让我们采用二阶导数的方法:

$$\frac{\mathrm{d}^2C}{\mathrm{d}r^2} = 4\pi\left(\frac{32}{r^3} + 2\right).$$

当 r 在区间 $(0, \infty)$ 上时, 二阶导数始终为正; 具体说, 当 $r = 2$ 时, 它为正, 所以在这里必有局部最小值.

剩下需要做的, 只是找出当 $r = 2$ 时其他变量的值并写出结论. 确实, 当 $r = 2$ 时, 可以看到 $h = 16/r^2 = 4$ 和 $C = 4\pi rh + 4\pi r^2 = 48\pi$. 这意味着成本最低的形状是半径为 2 英寸、高度为 4 英寸的圆柱体, 这时每个罐子的成本为 48π 美分, 也就是大约 1.50 美元. (这对一个普通罐子来说是相当昂贵了!) 注意到在这种情况下, 罐子的直径和高度相同.

接下来解答问题的第二部分. 这里其他条件不变, 只是现在要多加 14 美分每

英寸的焊接成本, 所以成本 C 的表达式会发生变化. 那么焊接每个罐子要花多少钱呢? 我们需要焊接盖和底, 所以面对的是一个圆的周长的两倍. 这意味着每个罐子需要焊接 $2\pi r$ 英寸的两倍, 也就是 $4\pi r$ 英寸. 这导致每个罐子成本增加 $14\times 4\pi r$ 美分, 所以 C 的新表达式为

$$C = 4\pi\left(\frac{16}{r} + r^2\right) + 14\times 4\pi r = 4\pi\left(\frac{16}{r} + r^2 + 14r\right).$$

(提出那个讨厌的 4π 是个好主意.) 不管怎样, 现在两边求导可得

$$\frac{\mathrm{d}C}{\mathrm{d}r} = 4\pi\left(-\frac{16}{r^2} + 2r + 14\right),$$

而当

$$-\frac{16}{r^2} + 2r + 14 = 0$$

时它为 0. 为了求解上述方程, 两边同时乘以 r^2 再除以 2, 得到

$$r^3 + 7r^2 - 8 = 0.$$

(确保你检验过这个结果是对!) 太好了! 现在我们有个三次方程要解. 幸运的是, 简单如 $r=1$ 就正好能用. 所以可通过做长除法看出另一个因子为 (r^2+8r+8). (检验一下!) 于是有

$$(r-1)\left(r^2+8r+8\right) = 0,$$

这意味这 $r=1$ 或 $r^2+8r+8=0$. 后者的解为 $\dfrac{-8\pm\sqrt{32}}{2}$, 又由于 $\sqrt{32}$ 的值约为 6, 这两个解都为负. 所以唯一 r 为正的临界点是 $r=1$. 再一次地, 由于端点处的成本为无穷大 (原因同前, 而算上焊接显然不会使其更便宜), 所以这个点为最小值. 作为检验, 我们有

$$\frac{\mathrm{d}^2C}{\mathrm{d}r^2} = 4\pi\left(\frac{32}{r^3} + 2\right),$$

它与之前的实际上是一样的. 所以它为正, 原始函数图像凹向上, 而在 $r=1$ 确实有最小值.

现在我们只需做代入. 可以看到 $h = 16/r^2 = 16$, 以及 $C = 4\pi(16/1 + 1^2 + 14\times 1) = 124\pi$ 美分, 也就是大约 4 美元. 看来需要想办法降低成本了! 不管怎样, 现在理想的罐子形状是半径为 1 英寸、高度为 16 英寸的圆柱体, 此时每个罐子的成本为 124π 美分. 注意到现在的最优半径比我们在问题第一部分计算出来的要小, 这是说得通的, 毕竟更小的半径降低了昂贵的焊接成本.

13.1.5　在最优化问题中使用隐函数求导



在我们转入最后一个例子之前, 先重新看一下上节问题的第一部分. 当时我们知道

$$C = 4\pi rh + 4\pi r^2 \quad 和 \quad r^2h = 16,$$

并通过消去 h 而最小化 C. 但还有一种最小化 C 的方法, 那就是在两边同时关于 r 作隐函数求导, 毕竟 r 是我们想保留的变量. (关于隐函数求导, 可回顾一下 8.1 节.) 这样得到

$$\frac{\mathrm{d}C}{\mathrm{d}r} = 4\pi\left(h + r\frac{\mathrm{d}h}{\mathrm{d}r} + 2r\right) \quad 和 \quad 2rh + r^2\frac{\mathrm{d}h}{\mathrm{d}r} = 0.$$



检验一下以确信计算是正确的. 不管怎样, 如果从第二个方程求解 $\mathrm{d}h/\mathrm{d}r$, 由于 $r \neq 0$, 就有

$$\frac{\mathrm{d}h}{\mathrm{d}r} = -\frac{2rh}{r^2} = -\frac{2h}{r}.$$

把它代入第一个方程, 得到

$$\frac{\mathrm{d}C}{\mathrm{d}r} = 4\pi\left(h + r\times\left(-\frac{2h}{r}\right) + 2r\right) = 4\pi(h - 2h + 2r) = 4\pi(2r - h).$$

所以当 $2r = h$ 时, $\mathrm{d}C/\mathrm{d}r = 0$, 而这正是我们之前得到的结果! 为了证明这个临界点是最小值, 对上式关于 r 再次求导, 得到

$$\frac{\mathrm{d}^2C}{\mathrm{d}r^2} = 4\pi\left(2 - \frac{\mathrm{d}h}{\mathrm{d}r}\right) = 4\pi\left(2 + \frac{2h}{r}\right).$$



(这里用到了刚才得到的 $\mathrm{d}h/\mathrm{d}r = -2h/r$.) 注意到上式的右边始终为正, 所以 C 关于 r 的图像是凹向上的, 而确实是有一个最小值. 当然, 知道当 $2r = h$ 时有最小值, 这并不没有告诉我们各个变量实际上是多少. 为此, 只需把 $2r = h$ 代入 $r^2h = 16$ 得到 $2r^3 = 16$, 我们就再次有 r =2 和 h =4.

接下来, 看你是否能够自己通过隐函数求导去解决问题的第二部分, 要确保你得到了与我们之前相同的结果.

13.1.6　一个较难的最优化例子



假设距离海岸灯塔正东 8 英里处的海中有一处石油钻井平台. 该平台的备用发电机位于灯塔正北 2 英里处. 你需要在发电机与平台之间铺设海底电缆. 在海岸以东 1 英里范围以内的海水较浅, 但随后海水急剧变深. 工作人员在浅海区铺设 1 英里电缆只需 1 天时间, 但在深海区铺设 1 英里电缆需要 5 天时间. 证明以图 13-7 所示方式 (所有单位都为英里) 铺设电缆是最快的, 并求出在这种情况下电缆的长度.

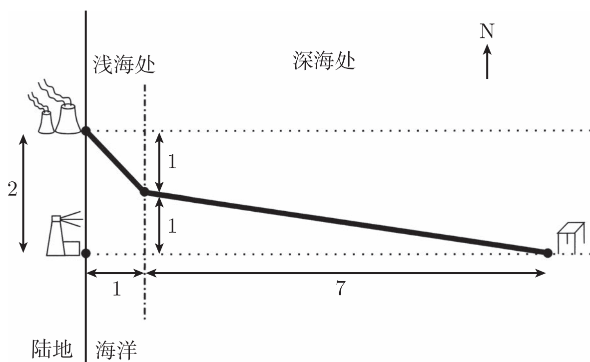


图 13-7

嗯, 这个问题看上去很难. 首先, 我们注意到上图至少是贴近现实的. 疯子才会把电缆铺设得绕来绕去, 因为这只会增加它的长度. 另一方面, 我们需要小心确定电缆应该从哪里从浅海区进入深海区. 一旦转换点确定了, 从发电机沿直线铺设电缆到这一点, 再从这一点沿直线铺设电缆到平台就显得聪明多了. 再一次地, 把转换点确定在发电机以北或平台以南也是疯狂之举 —— 这只会白白增加所花时间. 图 13-8 给出了一些合理的可能方案.

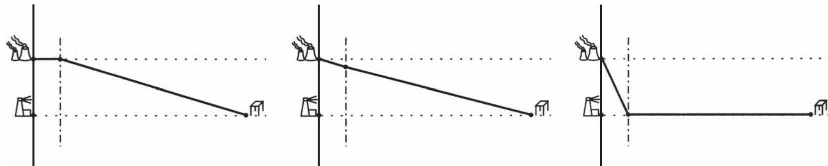


图 13-8

在第一个图中, 有太多的电缆位于深海区, 所以这很可能不是个好主意. 第二个图展示了所用电缆最少的情形, 但这并不意味着所花时间最少：仍有相当多电缆位于深海区. 第三个图展示了深海区电缆最少的情形, 但这样做的代价是使大量电缆位于浅海区. 这几次试探再次确认了, 最快的解决方案很可能是介于第二个图和第三个图所示的情形.

现在是时候引入一些变量了. 令 y, z, s 和 t 的含义如图 13-9 所示.

也就是说, s 是电缆在浅海区的长度, t 则是其在深海区的长度;[①] 此外, y 是转换点下距灯塔与平台的水平连线之间的距离, z 则是转换点上距发电机的水平延长线之间的距离, 所以 $y + z =2$. 我们想要证明, 最快的电缆铺设方式是当 y 和 z 都为 1 时. 我们已经看到 y 和 z 应该位于区间 $[0, 2]$, 但实际上我们甚至不需要假设这一点.

① 我猜电缆在深海区的长度本该称为 d, 但 $\mathrm{d}d/\mathrm{d}x$ 不是看起来很奇怪吗? 所以在微积分中不要把 d 作为变量来用!

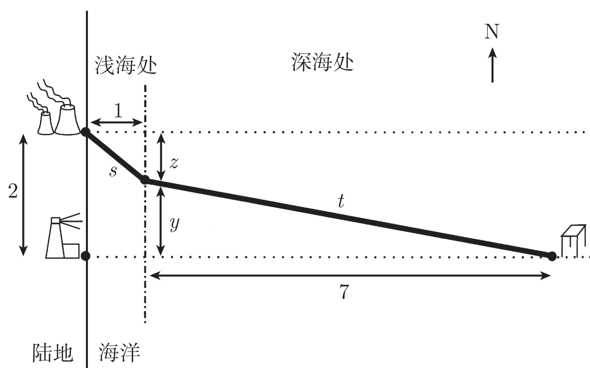


图　13-9

我们还想要求出这时铺设电缆所需的总时间. 由于在浅海区是 1 天每英里, 一共有 s 英里, 所以共需 $1 \times s = s$ 天去完成浅海区的作业. 类似地, 在深海区是 5 天每英里, 所以共需 $5t$ 天. 令 T 代表总天数, 有

$$T = s + 5t.$$

这就是我们想要最小化的量. 接下来, 需要找到关于 s 和 t 的方程. 为此, 使用勾股定理两次, 得到

$$s^2 = z^2 + 1,$$
$$t^2 = y^2 + 49.$$

对两个方程分别开根号, 并把结果代入 T 的表达式, 便有

$$T = \sqrt{z^2 + 1} + 5\sqrt{y^2 + 49}.$$

又由于 $y + z = 2$, 可用 $2 - y$ 来替代 z, 并得到



$$T = \sqrt{(2-y)^2 + 1} + 5\sqrt{y^2 + 49}.$$

我留给你来对上式求导, 并确认

$$\frac{\mathrm{d}T}{\mathrm{d}y} = -\frac{2-y}{\sqrt{(2-y)^2+1}} + \frac{5y}{\sqrt{y^2+49}}.$$

我们想要证明, 当 $y = 1$ 时, 所花时间最短. 让我们把这个值代入上式, 看能得到什么:

$$\frac{\mathrm{d}T}{\mathrm{d}y} = -\frac{1}{\sqrt{1^2+1}} + \frac{5}{\sqrt{1+49}} = -\frac{1}{\sqrt{2}} + \frac{5}{\sqrt{50}} = -\frac{1}{\sqrt{2}} + \frac{5}{5\sqrt{2}} = 0.$$

啊哈, $y = 1$ 是一个临界点! 所以至少存在可能, 它是全局最小值. 不过, 我们仍然需要证明这一点. 方法之一是求二阶导数. 经过相当一番折腾后, 你可以表明

$$\frac{\mathrm{d}^2T}{\mathrm{d}y^2} = \frac{1}{((2-y)^2+1)^{3/2}} + \frac{245}{(y^2+49)^{3/2}}.$$

二阶导数始终为正, 所以图像是凹向上的, 而 $y=1$ 确实为局部最小值. 事实上, 它必定为唯一的局部最小值! 确实, 如果还存在其他临界点, 那么它们也必都为局部最小值, 因为二阶导数始终为正. 但不可能有一大堆局部最小值, 却不让它们之间出现局部最大值, 所以其实并没有更多的临界点. 这意味着 $y=1$ 也是全局最小值, 而这正是我们想要的.

我们就快完成求解了：接下来只需把 $y=1$ 代入 T 的方程, 得到

$$T = \sqrt{(2-1)^2+1} + 5\sqrt{1^2+49} = \sqrt{2} + 5\sqrt{50} = \sqrt{2} + 25\sqrt{2} = 26\sqrt{2}.$$



所以总共需要 $26\sqrt{2}$ 天, 也就是大约 36.75 天.

在我们转入下一个话题之前, 再快速看一下另一种证明 $y=1$ 是最小值的方法. 这里的技巧是, 将表达式

$$\frac{\mathrm{d}T}{\mathrm{d}y} = -\frac{2-y}{\sqrt{(2-y)^2+1}} + \frac{5y}{\sqrt{y^2+49}}$$

巧妙加以重写. 对右边第一项, 分子分母同时除以 $(2-y)$, 而对第二项, 同时除以 y. 再做一个合理的假设, y 和 $2-y$ 都为正, 则有

$$\frac{\mathrm{d}T}{\mathrm{d}y} = -\frac{1}{\sqrt{1+\dfrac{1}{(2-y)^2}}} + \frac{5}{\sqrt{1+\dfrac{49}{y^2}}}.$$

随着 y 越来越大, 会发生什么？好吧, $(2-y)$ 会越来越小, $(2-y)^2$ 也是如此, 所以 $1/(2-y)^2$ 会越来越大. 这意味着第一项的分母会越来越大, 所以它的倒数会越来越小, 但倒数的相反数还是会越来越大. 这样细细想来, 可以得出结论：随着 y 越来越大, 第一项也会越来越大. 以同样的方法, 随着 y 越来越大, $49/y^2$ 会越来越小, 所以第二项的分母会越来越小, 但整个分式还是会越来越大.

这样我们较为轻松地证明了, $\mathrm{d}T/\mathrm{d}y$ 是增函数, 至少在 $(0, 2)$ 区间上是增函数. 由于 $\mathrm{d}T/\mathrm{d}y$ 是增函数, 所以它的导数 $\mathrm{d}^2T/\mathrm{d}y^2$ 为正! 这样, 无须实际计算出二阶导数就证明了它为正, 进而再一次地得到结论, $y=1$ 为最小值.

13.2 线 性 化



现在我们开始使用导数去估算特定的量. 例如, 假设想不借助计算器就得到 $\sqrt{11}$ 的一个较好估算. 我们知道 $\sqrt{11}$ 比 $\sqrt{9}=3$ 略大, 所以显然可以说 $\sqrt{11}$ 大约比 3 多一点. 这没问题, 但其实可以不费太多劲就做出一个好得多的估算. 下面是具体做法.

先设 $f(x)=\sqrt{x}$, $x\geqslant 0$. 我们想要估算 $f(11)=\sqrt{11}$ 的值, 因为不知道其确切值. 另一方面, 我们知道 $\sqrt{9}$ 确切是多少 —— 它就是 $\sqrt{9}=3$. 由于已知 $f(x)$ 当 $x=9$ 时的值, 不妨让我们绘出 $y=f(x)$ 的函数图像, 并画出一条通过点 (9, 3) 的切线, 就像图 13-10.

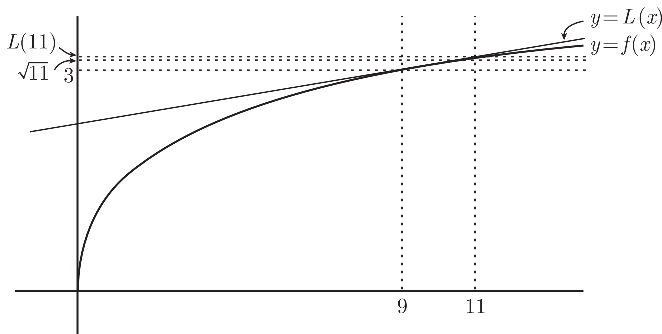


图　13-10

这条切线, 我标记为 $y=L(x)$, 在 $x=9$ 附近非常接近于曲线 $y=f(x)$. 当 $x=0$ 附近, 它就没有那么接近了. 但这无关紧要, 因为我们想要估算的是 $f(11)$, 而 11 是非常接近 9 的. 在上图中, 切线和曲线在 $x=11$ 处非常接近. 这意味着 $L(11)$ 是对 $f(11)=\sqrt{11}$ 的很好近似. 的确, 看看上图中这两个值在 y 轴上有多么接近吧!

不过, 如果不能实际计算出 $L(11)$, 那么刚才的一切就都是空话. 那么让我们来算吧. 线性函数 $L(x)$ 通过点 (9, 3), 并且由于它与曲线 $y=f(x)$ 在 $x=9$ 相切, 所以 $L(x)$ 的斜率为 $f'(9)$. 又由于 $f'(x)=1/(2\sqrt{x})$, 所以 $f'(9)=1/\left(2\sqrt{9}\right)=1/6$. 因此, $L(x)$ 斜率为 1/6, 并通过点 (9, 3). 于是其方程为

$$y-3=\frac{1}{6}(x-9),$$

化简可得 $y=x/6+3/2$. 也就是说,

$$L(x)=\frac{x}{6}+\frac{3}{2}.$$

现在, 只需将 $x=11$ 代入上式, 算得 $L(11)$ 的值：

$$L(11)=\frac{11}{6}+\frac{3}{2}=\frac{10}{3}=3\frac{1}{3}.$$

因此, 我们得到结论：

$$\sqrt{11}\approx 3\frac{1}{3}.$$

这可比之前的 3 多一点要好得多! 事实上, 可以使用计算器算得 $\sqrt{11}$ 约为 3.317 (精确到第三位小数), 所以我们的近似值 $3\frac{1}{3}$ 还是相当不错的.

13.2.1 线性化问题：一般方法

让我们将上述例子中所用方法一般化. 如果你想要估算某个量, 首先试着把它写成某个适当的函数 $f(x)$ 的值. 在上述例子中, 我们想要估算 $\sqrt{11}$, 所以设函数 $f(x)=\sqrt{x}$, 并意识到我们感兴趣的是 $f(11)$ 的值.

接下来, 我们选某个与 x 很接近的数 a, 并使得 $f(a)$ 容易计算. 在这个例子中, 我们无法处理 $f(11)$, 但容易计算 $f(9)$, 因为 9 开根号很容易. 我们也可以选择 $a=$25, 毕竟 25 开根号也很容易, 但这就不如选 9 好, 因为 25 离 11 相当远了.

再次, 已知函数 f 和特殊值 a, 我们找出通过曲线 $y=f(x)$ 上点 $(a, f(a))$ 的切线. 这条切线的斜率为 $f'(a)$, 所以其方程为

$$y-f(a)=f'(a)(x-a).$$

如果设切线为 $y=L(x)$, 则在上述方程两边同时加上 $f(a)$, 得到

$$L(x)=f(a)+f'(a)(x-a).$$

这个线性函数 L 被称为 f 在 $x=a$ 处的**线性化**. 回想一下, 我们将把 $L(x)$ 作为 $f(x)$ 的近似. 所以有

$$f(x)\approx L(x)=f(a)+f'(a)(x-a),$$

并知道当 x 很接近于 a 时, 这个近似是非常好的! 事实上, 当 x 实际上等于 a 时, 这个近似是完美的! 此时上述方程的两边都为 $f(a)$. 不过, 这并没什么用, 毕竟对 $f(a)$ 我们已经知根知底了. 这样, 现在有了对 $f(x)$ 在 x **接近于** a 时的近似.

让我们用上一节的例子来检验一下公式是否有效. 我们有 $f(x)=\sqrt{x}$ 和 $a=9$. 显然 $f(a)=f(9)=3$; 又由于 $f'(x)=1/\left(2\sqrt{x}\right)$, 我们有 $f'(9)=1/\left(2\sqrt{9}\right)=1/6$. 根据上述公式, f 的线性化为

$$L(x)=f(a)+f'(a)(x-a)=3+\frac{1}{6}(x-9).$$

这与之前得到的 $L(x)=\dfrac{x}{6}+\dfrac{3}{2}$一致, 我们当时正是用它求得 $\sqrt{11}\approx 3\dfrac{1}{3}$. 现在, 你知道怎样估算 $\sqrt{8}$ 吗? 注意到 8 也接近于 9, 所以可以使用同一个线性化：



$$\sqrt{8}=f(8)\approx L(8)=3+\frac{1}{6}(8-9)=\frac{17}{6}.$$

因此, 公式 $L(x)=3+\dfrac{1}{6}(x-9)$ 给出了所有 x 接近于 9 的 $\sqrt{x}$ 的很好近似, 而不单单是 11.

另一方面, 假设你还想要估算 $\sqrt{62}$. 这时使用 $L(62)$ 作为近似就不是很理想了. 让我们看看要是这样做的话会发生什么：



$$L(62)=3+\frac{62-9}{6}=11\frac{5}{6}.$$

等一下, $\sqrt{62}$ 本该比 $\sqrt{64}=8$ 稍小. 但 $L(62)$ 的值, 也就是 $11\frac{5}{6}$, 却比 8 大多了. 这里的问题在于, 线性化是在 $x=9$ 这点做的, 而 62 离 9 太远了, 所以这个近似就不太好了. 为了估算 $\sqrt{62}$, 更合适的做法是使用在 $x=64$ 处的线性化. 因此, 设 $a=64$, 我们有 $f(a)=8$ 和 $f'(a)=1/\left(2\sqrt{64}\right)=1/16$. 这意味着新的线性化为

$$L(x)=f(a)+f'(a)(x-a)=8+\frac{1}{16}(x-64).$$

当 $x=62$ 时, 有

$$\sqrt{62}=f(62)\approx L(62)=8+\frac{1}{16}(62-64)=7\frac{7}{8}.$$

这个近似就比 $11\frac{5}{6}$ 说得通多了.

13.2.2 微分

再来看一下刚才的一般方法. 我们看到

$$f(x)\approx f(a)+f'(a)(x-a).$$

不妨定义 $\Delta x=x-a$, 这样 $x=a+\Delta x$. 上述公式则变为

$$\boxed{f(a+\Delta x)\approx f(a)+f'(a)\Delta x.}$$

这时的情形可用图 13-11 表示.

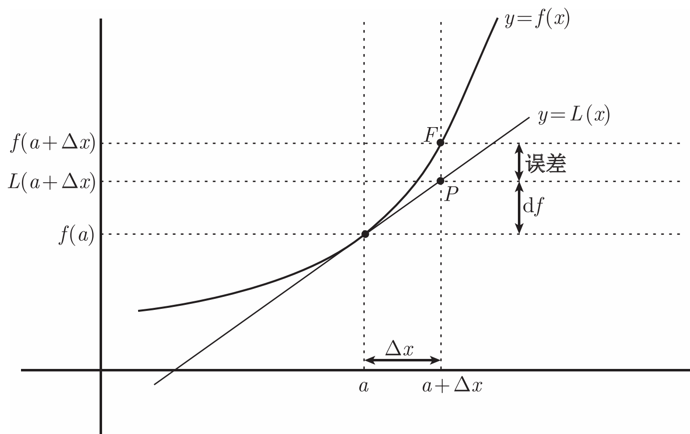


图 13-11

图中显示了曲线 $y=f(x)$ 以及线性化 $y=L(x)$, 后者是前者在 $x=a$ 处的切线. 我们想要估算 $f(a+\Delta x)$ 的值, 也就是图中点 F 的高度. 但作为近似, 我们实际上使用的是 $L(a+\Delta x)$, 也就是图中点 P 的高度. 这两个量之间的差被称为 “误差”, 我们会在 13.2.4 节再仔细讨论这个问题.

在上图中, 还有个量被标记了出来, 那就是 $\mathrm{d}f$, 也就是点 P 和 $f(a)$ 的高度之差. 我们需要把它加到 $f(a)$ 上才能得到估算. 由于 $L(a+\Delta x)=f(a)+f'(a)\Delta x$, 有

$$\boxed{\mathrm{d}f=f'(a)\Delta x.}$$

量 $\mathrm{d}f$ 被称为 f 在 $x=a$ 处的**微分**. 它是对当 x 从 a 变化为 $a+\Delta x$ 时 f 的变化量的近似.

我们其实在前面已经遇到过类似情况. 在 5.2.7 节, 如果 $y=f(x)$, 则有

$$f'(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}.$$

这意味着 x 的微小变化会引起 y 的变化, 而后者的变化量约为前者的 $f'(x)$ 倍. 这也正是 $\mathrm{d}f=f'(a)\Delta x$ 所说的, 只是这时的变化是从 $x=a$ 起.

例如, 假设想要估算 $(6.01)^2$. 设 $f(x)=x^2$ 且 $a=6$, 则轻松可得 $f'(x)=2x$, 所以 $f'(6)=12$. 我们想要知道当 x 从 6 起增加 0.01 时 $f(x)$ 会发生什么变化, 所以应该设 $\Delta x=0.01$. 于是有

$$\mathrm{d}f=f'(a)\Delta x=f'(6)(0.01)=12\times(0.01)=0.12.$$

因此, 如果把 0.12 加到 $f(a)$ 的值上, 应该能得到一个不错的近似. 由于 $f(a)=f(6)=6^2=36$, 这意味着 $(6.01)^2\approx 36.12$. 现在让我们再次回过头去看一下 5.2.7 节, 在那里解答了相同的例题, 使用了基本相同的方法 —— 只是现在有了更好用的公式, 仅此而已.

下面是另一个展示怎样使用微分的例子. 假设用一把尺子测得一个圆球的直径为 6 英寸, 但这个测量结果有正负 0.5% 的误差. 如果使用这个测量结果去计算该球的体积, 所得结论的准确度有多高? 让我们使用微分来求解, 至少是近似地. 如果球的半径为 r、直径为 D、体积为 V, 则 $r=D/2$, 有

$$V=\frac{4}{3}\pi r^3=\frac{4}{3}\pi\left(\frac{D}{2}\right)^3=\frac{\pi D^3}{6}.$$

当 $D=6$ 时, 有 $V=\pi(6)^3/6=36\pi$. 所以算得该球的体积为 36π 立方英寸, 但其真值可能会比这略大或略小. 为了找出究竟多多少或少多少, 让我们使用前边的加框公式 $\mathrm{d}f=f'(a)\Delta x$. 这里需要用 V 代替 f, 用 6 代替 a, 以及用 D 代替 x, 以得到相应公式:

$$\mathrm{d}V=V'(6)\Delta D.$$

对前面 V 的表达式两边关于 D 求导, 得到

$$V'(D)=\frac{\pi(3D^2)}{6}=\frac{\pi D^2}{2}.$$

这意味着 $V'(6)=18\pi$, 所以

$$\mathrm{d}V = 18\pi\Delta D.$$

这个方程意味着, 如果将直径 D 从 6 变为 $6+\Delta D$, 那么体积 V 会改变大约 $18\pi\Delta D$. 在我们的例子中, 直径的真值可能比 6 多或少 0.5%, 也就是 $0.005\times 6 = 0.03$ 英寸. 所以 ΔD 最高可能到 0.03, 或最低可能到 -0.03. 在这种最糟糕的情况下, 我们有

$$\mathrm{d}V = 18\pi\times(\pm 0.03) = \pm 0.54\pi.$$

这是对测量引起的误差的一个很好估算, 所以可以说该球的体积为 36π 立方英寸, 误差约 0.54π 立方英寸. 由于原始问题将直径的测量误差用百分比来表示, 所以很可能也应该这样来表示体积的误差. 近似误差 $\mathrm{d}V = \pm 0.54\pi$ 占体积 $V = 36\pi$ 的百分比为

$$\frac{\mathrm{d}V}{V}\times 100\% = \frac{\pm 0.54\pi}{36\pi}\times 100\% = \pm 1.5\%.$$

换句话说, 体积测量的相对误差大约是原始直径测量的相对误差的三倍. 当使用一个一维的度量去计算一个三维的量时, 原始的测量误差就会复合积累到这个程度.

13.2.3　线性化的总结和例子

以下是估算或近似计算一个难搞定的数的基本策略.

(1) 写出主要公式：

$$\boxed{f(x)\approx L(x) = f(a) + f'(a)(x-a).}$$

(2) 选择一个函数 f 以及一个数 x, 使得这个难搞定的数等于 $f(x)$. 另外, 选取一个接近于 x 的 a, 并使得 $f(a)$ 可以容易算得.

(3) 对 f 求导, 找出 $f'(x)$.

(4) 在上述方框公式中, 用实际的函数分别替代 f 和 f', 用你选定的实际数值替代 a.

(5) 最后, 把第二步中的 x 值代入公式加以计算. 另外注意到微分 $\mathrm{d}f$ 等于量 $f'(a)(x-a)$.



下面让我们来看几个例子. 首先, 你会怎样估算 $\sin(11\pi/30)$? 先写出标准公式：

$$f(x)\approx L(x) = f(a) + f'(a)(x-a).$$

我们需要求的是某数的正弦值, 所以设 $f(x) = \sin x$. 并且我们对当 $x = 11\pi/30$ 时的函数值感兴趣. 接下来需要选择一个接近于 $11\pi/30$ 的数 a, 并使得 $f(a)$ 容易计算. 当然, $f(a)$ 就是 $\sin(a)$. 那么什么数既接近于 $11\pi/30$, 其正弦值又容易计算

呢? $10\pi/30$ 怎么样? 毕竟它就是 $\pi/3$, 而我们显然很清楚 $\sin(\pi/3)$ 的值. 所以不妨设 $a=\pi/3$.

这样就完成了前两步, 下面进入第 (3) 步. 求导得到 $f'(x)=\cos x$, 所以线性化公式变为

$$f(x)\approx L(x)=\sin\left(\frac{\pi}{3}\right)+\cos\left(\frac{\pi}{3}\right)\left(x-\frac{\pi}{3}\right).$$

由于 $f(x)=\sin x$, 上式可化简为

$$\sin(x)\approx L(x)=\frac{\sqrt{3}}{2}+\frac{1}{2}\left(x-\frac{\pi}{3}\right).$$

最后, 代入 $x=11\pi/30$, 得到

$$\sin\left(\frac{11\pi}{30}\right)\approx L\left(\frac{11\pi}{30}\right)=\frac{\sqrt{3}}{2}+\frac{1}{2}\left(\frac{11\pi}{30}-\frac{\pi}{3}\right)=\frac{\sqrt{3}}{2}+\frac{\pi}{60}.$$

这可能看起来仍然很糟, 但至少这个估算没有涉及三角函数 —— 有的只是 π 和 $\sqrt{3}$, 而这两个数并不是很难处理.

现在, 考虑以下例子：用线性化找到 $\ln(0.99)$ 的一个近似值. 这次设 $f(x)=\ln(x)$, 并且感兴趣的是当 $x=0.99$ 时 $f(x)$ 的值. 接近于 0.99 并就取对数而言容易算得的数是 1, 所以设 $a=1$. 由于 $f(x)=\ln(x)$ 且 $f'(x)=1/x$, 公式 $f(x)\approx L(x)=f(a)+f'(a)(x-a)$ 变为

$$\ln(x)\approx L(x)=\ln(1)+\frac{1}{1}(x-1).$$

又由于 $\ln(1)=0$, 有

$$\ln(x)\approx x-1.$$

代入 $x=0.99$, 得到

$$\ln(0.99)\approx L(0.99)=0.99-1=-0.01.$$

这样, 我们就完成了.

更一般地, 你会怎样估算当 h 为任意很小的数时 $\ln(1+h)$ 的值? 事实上, 可以使用刚刚找到的线性化 $f(x)\approx L(x)=x-1$, 去估算 $\ln(1+h)$. 只需用 $1+h$ 替代 x, 便可见 $\ln(1+h)\approx L(1+h)=(1+h)-1$. 也就是说, 当 h 很小时,

$$\ln(1+h)\approx h.$$

这其实并不该让人感到意外. 在 9.4.3 节中, 我们已经看到过

$$\lim_{h\to 0}\frac{\ln(1+h)}{h}=1.$$

所以早已知道, 当 h 是个很小的数时, $\ln(1+h)$ 近似等于 h.

最后, 当 h 很小时, $\ln(\mathrm{e}+h)$ 的近似值又是多少呢? 现在, 我们需要一个不同的线性化, 因为量 $(\mathrm{e}+h)$ 接近于 e 而非 1. 所以设 $a=\mathrm{e}$, 并且再一次地用到

$f(x) = \ln(x)$ 和 $f'(x) = 1/x$. 于是有

$$f(x) \approx L(x) = f(a) + f'(a)(x - a) = \ln(\mathrm{e}) + \frac{1}{\mathrm{e}}(x - \mathrm{e}).$$

由于 ln(e)=1, 有

$$\ln(x) \approx L(x) = 1 + \frac{x}{\mathrm{e}} - 1 = \frac{x}{\mathrm{e}}.$$

当 $x = \mathrm{e} + h$ 时, 便有

$$\ln(\mathrm{e} + h) \approx L(\mathrm{e} + h) = \frac{\mathrm{e} + h}{\mathrm{e}} = 1 + \frac{h}{\mathrm{e}}.$$

也就是说, 当 h 很小时, $\ln(\mathrm{e}+h) \approx 1 + h/\mathrm{e}$. 这与上一例题中的结论是很不同的; 在上一例题中, 当 h 很小时, $\ln(1 + h) \approx h$. 所以一切都取决于 a 的值.

13.2.4　近似中的误差

我们一直在用 $L(x)$ 作为 $f(x)$ 的近似, 但它们并不是一回事. 那么我们用 $L(x)$ 代替 $f(x)$ 的做法错得有多离谱呢? 解答这个问题的方法是, 考虑这两个量之间的差. 它们的差越小, 近似就越精确. 所以, 设

$$r(x) = f(x) - L(x),$$

其中 $r(x)$ 是使用在 $x = a$ 处的线性化来估算 $f(x)$ 时的误差.[1] 结果表明, 如果函数 f 的二阶导数存在, 至少在 x 和 a 之间存在, 那么对于 $r(x)$ 就有一个很好的公式:[2]

$$r(x) = \frac{1}{2}f''(c)(x - a)^2, \quad \text{其中 } c \text{ 为在 } x \text{ 和 } a \text{ 之间的某个数}.$$

但问题是, 我们不知道 c 的值, 只知道它在 x 和 a 之间. 上述公式与 11.3 节讨论过的中值定理有关系. 该定理告诉了我们数 c 的性质, 却没有透露关于它的更多信息, 所以我们不该奇怪它也在这里出现了.

我们可从上述公式看出两件事. 首先, 注意到 $(x - a)^2$ 项始终为正. 这意味着 $r(x)$ 的符号与 $f''(c)$ 的符号相同. 因此, 如果知道函数图像是凹向上的, 至少在 x 和 a 之间是如此, 那么就知道 $r(x)$ 为正. 又由于 $r(x) = f(x) - L(x)$, 所以就有 $f(x) > L(x)$. 这意味着近似值比实际值偏小, 我们低估了. 前面图 13-11 所示的情形便是如此. 另一方面, 如果函数图像是凹向下的, 那么 $f''(c)$ 必定为负, 从而可知 $f(x) < L(x)$. 这意味着近似是高估了.

例如, 在 13.2 节开头部分估算 $\sqrt{11}$ 的例子中, 我们使用了 $f(x) = \sqrt{x}$. 通过算得 $f'(x) = 1/(2\sqrt{x})$ 和 $f''(x) = -1/(4x\sqrt{x})$, 可以看出函数图像始终是凹向下的. 或者也可以通过画函数图像看出来. 不管是哪种情况都可以看出, 得到的近似

---

① $r(x)$ 中的字母 r 代表“余数” (remainder), 因为它是当你除去线性化后余下的部分.
② 证明请参见附录 A 中的 A.6.9 节.

值 $3\frac{1}{3}$ 必定比实际值略大.



总结而言,

- 如果 f'' 在 a 和 x 之间为**正**, 则通过线性化得出的估算是**低估**.
- 如果 f'' 在 a 和 x 之间为**负**, 则通过线性化得出的估算是**高估**.

现在让我们再来看一下刚才的误差方程. 如果对上述方程的两边取绝对值, 可得

$$|\text{误差}| = \frac{1}{2}|f''(c)||x-a|^2.$$

假设我们知道当 t 在 x 和 a 之间变化时, $|f''(t)|$ 的最大值是某数 M. 那么尽管我们不知道 c 的具体值, 仍能知道 $|f''(c)| \leqslant M$, 于是有

$$|\text{误差}| \leqslant \frac{1}{2}M|x-a|^2,$$

其中 M 是当 t 在 x 和 a 之间变化时, $|f''(t)|$ 的最大值. 实际上, 在上述等式中要紧的因子不是 M, 而是 $|x-a|^2$. 你看, 当 x 接近于 a 时, 量 $|x-a|$ 已经很小, 而平方之后它就更小了. (例如, 当对 0.01 求平方时, 你将得到一个更小的数 0.0001.) 这意味着误差很小, 所以近似很好!

让我们看一下如何将此应用到刚才估算 $\sqrt{11}$ 的例子. 设 $f(x)=\sqrt{x}, f'(x)=1/\left(2\sqrt{x}\right)$, $f''(x)=-1/\left(4x\sqrt{x}\right)$. 还取 $a=9$, $x=11$. 现在的问题是, 当 t 在 9 和 11 之间变化时, $|f''(t)|$ 的值最大会是多少? 很显然,



$$|f''(t)| = \frac{1}{4t\sqrt{t}}.$$

等式右边是个关于 t 的减函数, 所以当 t 最小, 也就是 $t=9$ 时, 函数值越大. 所以 $M=|f''(9)|$, 也就是 1/108. 于是可以得出结论:

$$|\text{误差}| \leqslant \frac{1}{2}M|x-a|^2 = \frac{1}{2}\frac{1}{108}|11-9|^2 = \frac{1}{54}.$$

因此, 以前我们说 $\sqrt{11} \approx 3\frac{1}{3}$, 现在则可以信心十足地说, 近似相当接近. 事实上, 在实际值的 $\pm 1/54$ 之内. 更准确地说, 有

$$3\frac{1}{3} - \frac{1}{54} \leqslant \sqrt{11} \leqslant 3\frac{1}{3} + \frac{1}{54}.$$

事实上, 由于我们已经知道 $3\frac{1}{3}$ 比 $\sqrt{11}$ 的实际值略大, 所以可以进一步说:

$$3\frac{1}{3} - \frac{1}{54} \leqslant \sqrt{11} \leqslant 3\frac{1}{3}.$$

现在, 再试一下 13.2.3 节所见的估算 $\ln(0.99)$ 的例子. 在那里, 得到 $\ln(0.99) \approx -0.01$. 那么这个近似到底有多好呢? 设 $f(x)=\ln(x)$, 有 $f'(x)=1/x$, $f''(x)=-1/x^2$. 由于二阶导数为负, 再一次地, 我们的估算偏大. 现在, 当 t 在 $a=1$ 和 $x=0.99$



之间变化时, $|f''(t)| = 1/t^2$ 的值最大会是多少? 再一次地, 等式右边为 t 的减函数, 所以最大值当 t =0.99 时取到. 于是有 $M = 1/(0.99)^2$, 并且估算误差为

$$|\text{误差}| \leqslant \frac{1}{2}M|x-a|^2 = \frac{1}{2}\frac{1}{0.99^2}|0.99-1|^2 = \frac{1}{20\,000(0.99)^2}.$$

这化简后约为 0.000 051, 非常非常小的数. 这意味着 -0.01 是对 $\ln(0.99)$ 一个非常好的近似. 更准确地说, 我们证明了不等式

$$-0.01 - \frac{1}{20\,000(0.99)^2} \leqslant \ln(0.99) \leqslant -0.01 + \frac{1}{20\,000(0.99)^2}.$$

事实上, 由于我们知道 -0.01 偏大, 所以再一次地, 可以简化上述不等式的右边, 写成

$$-0.01 - \frac{1}{20\,000(0.99)^2} \leqslant \ln(0.99) \leqslant -0.01.$$

这样就把 $\ln(0.99)$ 的值缩小到了一个相当小的范围.

在后面第 24 章讨论泰勒级数时, 我们会再次回到估算和误差的话题. 到时我们将不仅用到一阶导数, 还会用到二阶以及更高阶导数去求得更精确的近似.

13.3　牛　顿　法

下面是线性化的另一个有用应用. 假设现在要解一个形为 $f(x) = 0$ 的方程, 但你死活都解不出来. 所以你退而求其次, 试着猜测该方程有一个解, 并把它记为 a. 这时的情形可能如图 13-12 所示.

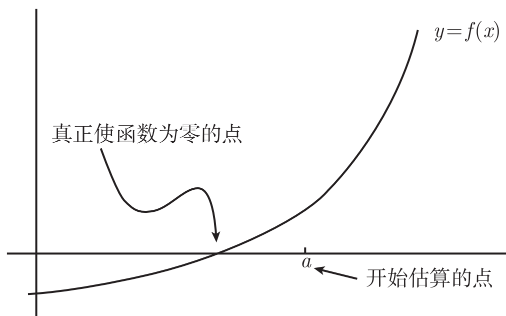


图　13-12

从图中可以看出, $f(a)$ 实际上并不等于零, 所以 a 其实并不是该方程的解, 它仅仅是解的一个近似或估算. 可以把它视为近似的第一次尝试, 所以在上图中把它标记为了 “初始的近似”. 牛顿法的基本思想是, 通过使用 f 在 $x = a$ 处的线性化来改善估算. (当然, 这意味着 f 需要在 $x = a$ 处是可导的.) 不管怎样, 让我们看一下图 13-13 的情形.

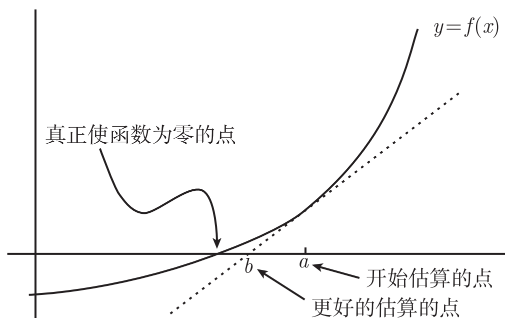


图 13-13

这个线性化的 x 轴截距记为 b, 并且显而易见, 相对于真正的零点, 它是个比 a 更好的近似. 这样从一个初始的猜测, 我们得到了一个更好的结果. 但 b 的值具体是多少呢? 好吧, 它就是线性化

$$L(x) = f(a) + f'(a)(x-a)$$

的 x 轴截距. 为了求 x 轴截距, 设 $L(x) = 0$, 则我们有 $f(a) + f'(a)(x-a) = 0$. 解得 x, 有

$$x = a - \frac{f(a)}{f'(a)}.$$

由于刚才把 x 轴截距记为 b, 于是有如下公式：

牛顿法 假设 a 是对方程 $f(x) = 0$ 的解的一个近似. 如果令

$$b = a - \frac{f(a)}{f'(a)},$$

则在很多情况下, b 是个比 a 更好的近似.

这并不是在所有情况下都成立, 所以我特意加上 “在很多情况下” 以防万一. 我们会在下一页再具体讨论这些特殊情况. 现在让我们先看一些例子. 假设

$$f(x) = x^5 + 2x - 1,$$

我们想求方程 $f(x) = 0$ 的解. 但首先该方程有解吗? 由于 f 是连续的, $f(0) = -1$(为负), $f(1) = 2$(为正), 根据介值定理 (参见 5.1.4 节), 该方程至少有一个解. 另一方面, $f'(x) = 5x^4 + 2$, 它始终为正, 所以 f 始终为增函数. 这意味着该方程至多有一个解 (参见 10.1.1 节). 这样就证明了该方程有唯一的解. 现在让我们从 0 开始逼近方程的解. 我们知道 $f(0) = -1$, 这并不是很接近于 0. 但没关系, 就使用牛顿法从 a =0 开始：

$$b = a - \frac{f(a)}{f'(a)} = 0 - \frac{f(0)}{f'(0)} = 0 - \frac{0^5 + 2(0) - 1}{5(0)^4 + 2} = \frac{1}{2}.$$

所以 $b=1/2$ 是个比 0 更好的近似. 确实, 可以算得 $f(1/2)=1/32$, 它相当接近于 0. 那么为什么不重复使用这个方法, 从而得到一个还要好的近似呢？当然可以! 这次取 $a=1/2$, 并再次算得：

$$b=a-\frac{f(a)}{f'(a)}=\frac{1}{2}-\frac{f(1/2)}{f'(1/2)}=\frac{1}{2}-\frac{1/32}{37/16}=\frac{18}{37}.$$

(这里用到了 $f'(1/2)=5\times(1/2)^4+2=37/16$.) 不管怎样, 这意味着 18/37 是个还要好的近似. 计算 $f(18/37)$, 会算得结果约为 0.0002, 这已经是非常非常小的数了. 数 18/37 确实是对 f 的真正零点的一个相当好的近似.

像这样重复使用 a 和 b 可能会引起混乱. 一种规避的方法是, 用 x_0 标记初始的猜测, 用 x_1 标记第一次改善, 用 x_2 标记基于 x_1 的第二次改善, 如此等等. 这样公式变为

$$x_1=x_0-\frac{f(x_0)}{f'(x_0)},\ x_2=x_1-\frac{f(x_1)}{f'(x_1)},\ x_3=x_2-\frac{f(x_2)}{f'(x_2)},\quad \text{等等}.$$

下面是另一个例子：求方程 $x=\cos x$ 的近似解. 首先设 $f(x)=x-\cos x$. 如果能估算出 f 的零点, 那么这个数也就是 $x=\cos x$ 的近似解. (5.1.4 节中已经使用过这个技巧了.) 让我们作个猜测 $x_0=\pi/2$, 则 $f(\pi/2)=\pi/2-\cos(\pi/2)=\pi/2$. 这是个相当不靠谱的猜测. 不过没关系, 由于 $f'(x)=1+\sin(x)$, 有 $f'(\pi/2)=1+\sin(\pi/2)=2$. 这意味着

$$x_1=x_0-\frac{f(x_0)}{f'(x_0)}=\frac{\pi}{2}-\frac{\pi/2}{2}=\frac{\pi}{4}.$$

所以 $x_1=\pi/4$ 是个更好的近似; 的确, $f(\pi/4)=\pi/4-1/\sqrt{2}$, 大约为 0.08. 现在重复使用这个过程. 由于 $f'(\pi/4)=1+\sin(\pi/4)=1+1/\sqrt{2}$,

$$x_2=x_1-\frac{f(x_1)}{f'(x_1)}=\frac{\pi}{4}-\frac{f(\pi/4)}{f'(\pi/4)}=\frac{\pi}{4}-\frac{\pi/4-1/\sqrt{2}}{1+1/\sqrt{2}}.$$

上式可化简为

$$x_2=\frac{1+\pi/4}{1+\sqrt{2}}=(1+\pi/4)(\sqrt{2}-1),$$

它实际上比 $\pi/4$ 稍小. 此外, 可算得 $f(x_2)$ 约为 0.0008. 这意味着 $x_2-\cos(x_2)$ 约为 0.0008, 所以数 x_2 是对方程 $x=\cos(x)$ 的解的一个相当好的近似. 当然, 可以重复使用这个方法, 以得到一个还要好的近似值 x_3, 但这时的计算就变得很糟糕了. 不过, 计算机和计算器倒是善于此道, 并且实际上它们常常使用牛顿法以给出很好的近似. (别忘了, 计算器给出的也**仅仅**是近似值! 即便小数点后有 10 位或 12 位, 它仍然不是确切的值, 尽管在大多数情况下这已经够用了.)



正如我们之前注意到的 (但没有给出解释), 有时牛顿法也会不起作用. 下面是失效的四种不同情况.

(1) **$f'(a)$ 的值接近于 0.** 显然, 如果

$$b = a - \frac{f(a)}{f'(a)},$$

则 $f'(a)$ 不能为 0, 否则 b 是没有定义的. 在这种情况下, 在 $x = a$ 处的切线不可能与 x 轴相交, 因为它是水平的! 即使 $f'(a)$ 很接近但不等于 0, 牛顿法仍会给出一个很糟糕的结果. 例如, 图 13-14 所示的情形.

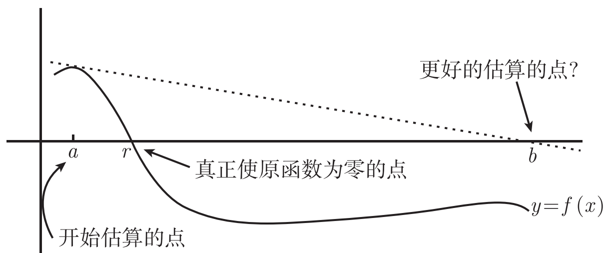


图 13-14

即便从一个相当好的近似 a 开始, 牛顿法给出的结果 b 还是远离真正的零点 r. 所以根本没有得到一个更好的近似. 为了避免出现这种情况, 要确保你的初始猜测不在函数 f 的临界点附近.

(2) **如果 $f(x) = 0$ 有不止一个解, 可能得到的不是你想要的那个解.** 例如在图 13-15 中, 如果你想估算左边的根 r, 并且猜测从 a 开始, 那么最终你估算的其实是另一个根 s.

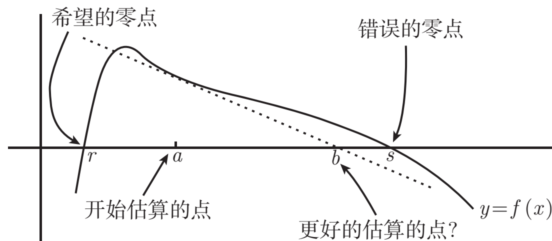


图 13-15

所以你应该稍微花些工夫, 选取一个接近于你想要的那个零点的初始猜测 a, 除非你确定只有一个解.

(3) **近似可能变得越来越糟.** 例如, 如果 $f(x) = x^{1/3}$, 方程 $f(x) = 0$ 唯一的解是 x =0. 如果你尝试对此使用牛顿法 (出于某种只有你自己知道的原因), 那么怪事就会出现. 你看, 除非从 $a = 0$ 开始, 否则会得到



$$b = a - \frac{f(a)}{f'(a)} = a - \frac{a^{1/3}}{a^{-2/3}/3} = -2a.$$

所以下一个近似值总是你初始值的 -2 倍. 例如, 如果从 $a = 1$ 开始, 那么下一个

近似值将是 -2. 如果重复这个过程, 将得到 4, -8, 16, 等等. 结果是离正确值 0 越来越远. 如果遇到这类情况, 牛顿法就无能为力了.

(4) **你可能陷入一个循环而无法自拔.** 有可能出现, 你通过估算 a 得到 b, 估算 b 却又得到 a. 重复这个过程是没有意义的, 因为你只是在兜圈子! 这种情况可能如图 13-16 所示.

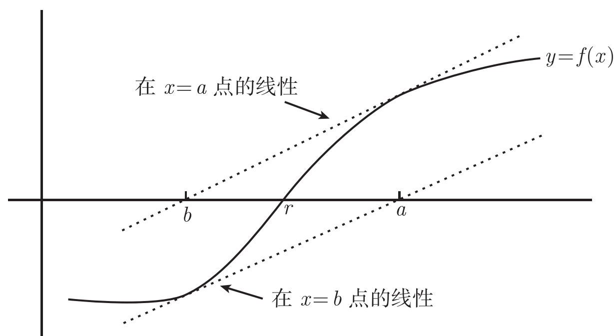


图　13-16

在 $x = a$ 处的线性化有 x 轴截距 b, 而在 $x = b$ 处的线性化有 x 轴截距 a, 所以牛顿法在这里就不灵了. 一个具体 (但有点复杂) 的例子是

$$f(x) = \left(x^2 - \frac{4+3\pi}{4-\pi}\right)\tan^{-1}(x).$$

如果从 $a = 1$ 开始, 我留给你来算出 $b = -1$. 由于 f 为奇函数, 显然再从 -1 开始的计算会再次得到 1. 真不幸我们陷入了一个循环! 不妨再选其他的初始猜测试试看. (顺便说一下, 对此类循环的研究引出了一类好看的分形, 你可能在某人的计算机屏保上看到过 ······)

第 14 章　洛必达法则及极限问题总结

在讲解导数的时候, 我们利用极限给导数下了基本定义. 现在, 我们要倒过来, 利用导数的知识求极限的值, 这种方法叫**洛必达法则**. 在介绍这个法则的各种类型及使用方法后, 我们会对目前已使用的计算极限的所有方法进行总结. 我们将学到如下知识点：

- 洛必达法则及使用该法则的四种极限情况;
- 对计算极限的方法进行总结.

14.1　洛必达法则

我们学过的大部分极限都是以下情况之一：

$$\lim_{x\to a}\frac{f(x)}{g(x)},\ \lim_{x\to a}(f(x)-g(x)),\ \lim_{x\to a}f(x)g(x)\ 和\ \lim_{x\to a}f(x)^{g(x)}.$$

有时你可以利用函数的连续性直接用 a 来替代 x 进行计算, 但这种方法可能解决不了问题. 例如, 考虑下列极限：

$$\lim_{x\to 3}\frac{x^2-9}{x-3},\ \lim_{x\to 0}\left(\frac{1}{\sin(x)}-\frac{1}{x}\right),\ \lim_{x\to 0^+}x\ln(x)\ 和\ \lim_{x\to 0}(1+3\tan(x))^{1/x}.$$

在第一种情况中, 用 3 来替代 x 得到 0/0 型的不定式. 第二种极限是当 $x\to 0$ 时的两个无穷大的差. 实际上, 当 $x\to 0^+$ 时, 两个算式都趋于正无穷; 当 $x\to 0^-$ 时, 两个算式都趋于负无穷. 所以我们可以把这种形式总结为 $\pm(\infty-\infty)$. 第三种极限 (关于 $x\ln(x)$) 是 $0\times(-\infty)$ 类型, 请记住当 $x\to 0^+$ 时, $\ln(x)\to-\infty$. 最后, 第四种极限是 1^∞, 看起来也很难求. 但幸运的是, 我们可以使用洛必达法则求解这四种极限.

第一种类型是两个函数的比 $f(x)/g(x)$, 最适合用这一法则, 我们称它为 "类型 **A**". 接下来的两种类型为 $f(x)-g(x)$ 和 $f(x)g(x)$, 都可以直接化归为类型 **A**, 所以我们分别叫它们为类型 **B1** 和 **B2**. 最后, 我们把关于指数型函数 $f(x)^{g(x)}$ 的类型叫作类型 **C**, 该类型可以化归为类型 **B2**, 从而再化归为类型 **A**. 让我们先分别看看这些类型, 然后在 14.1.6 节中进行总结.

14.1.1　类型 A：0/0

考虑下述形式的极限：

$$\lim_{x\to a}\frac{f(x)}{g(x)},$$

f 和 g 都是很好的可导函数. 如果 $g(a) \neq 0$, 那情况就太棒了, 我们可以直接用 a 替代 x 来求极限 $f(a)/g(a)$ 的值. 如果 $g(a) = 0$, 但是 $f(a) \neq 0$, 这时在 $x = a$ 点有垂直渐近线, 上述极限为 ∞, $-\infty$ 或不存在. (参照图 4-1 中四种情况的图像, 有助于你理解.)

还有另外一种可能是 $f(a) = 0, g(a) = 0$. 也就是说, 该分式 $f(a)/g(a)$ 是 0/0 型的不定式. 我们见过的大多数极限都是这种类型. 事实上, 每一个导数都是这种形式! 毕竟,

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h},$$

如果把 $h = 0$ 代入分式, 就会得到 0/0 型. 所以, 我们主要研究 $f(a) = 0$ 和 $g(a) = 0$ 的情况.

基本思想是这样的：因为 f 和 g 是可导函数, 所以可以在 $x = a$ 点处对它们线性化. 像上一章一样, 当 x 趋于 a 点的时候, 我们有：

$$f(x) \approx f(a) + f'(a)(x-a) \quad \text{和} \quad g(x) \approx g(a) + g'(a)(x-a).$$

现在, 假设 $f(a)$ 和 $g(a)$ 都为 0, 这说明

$$f(x) \approx f'(a)(x-a) \quad \text{和} \quad g(x) \approx g'(a)(x-a).$$

如果你用 $f(x)$ 除以 $g(x)$, 假设 $x \neq a$, 则有

$$\frac{f(x)}{g(x)} \approx \frac{f'(a)(x-a)}{g'(a)(x-a)} = \frac{f'(a)}{g'(a)}.$$

x 越接近于 a, 这个估算就越接近真实值. 这样,[1]就有了洛必达法则的一种表达式：

$$\boxed{\text{如果 } f(a) = g(a) = 0, \text{ 那么 } \lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)},}$$

假设等式右端的极限存在. (实际上还有另一个条件, 当 x 趋于但不等于 a 时, $g'(x)$ 不为 0. 遇到这种情况时, 你真的很不幸!) $f(a)$ 和 $g(a)$ 都为 0, 这个前提真的很重要, 否则将不能使用这个法则.

让我们用本章开始的例子来演示怎样用这个法则解决问题：

$$\lim_{x\to 3} \frac{x^2-9}{x-3}.$$

注意, 如果把 $x = 3$ 代入原函数, 你会发现分子分母都为 0, 这说明我们可以使用洛必达法则. 你所需要做的, 是对分式的分子分母分别求导. 注意：请不要使用除法法则! 求解的过程为

$$\lim_{x\to 3} \frac{x^2-9}{x-3} \overset{\text{l'H}}{=} \lim_{x\to 3} \frac{2x}{1} = 6.$$

注意到上边的等式中会有个 “l'H” 符号, 其作用是说明我们是用洛必达法则来解决问题的 (洛必达法则的英文是 l'Hôpital's Rule). 顺便说一下, 这道题其实也可以不

① 实际上, 我们还没有证明洛必达法则. 真正的证明请参阅附录 A 的 A.6.11 节.

使用洛必达法则, 因为分子 x^2-9 可以被因式分解为 $(x+3)(x-3)$, 所以计算过程可以是

$$\lim_{x\to 3}\frac{x^2-9}{x-3}=\lim_{x\to 3}\frac{(x-3)(x+3)}{x-3}=\lim_{x\to 3}(x+3)=3+3=6.$$

我们得到了同样的答案! 这说明计算是正确的.

这儿有一个更难的例子, 使用因式分解不能解决问题了：

$$\lim_{x\to 0}\frac{x-\sin(x)}{x^3}.$$

如果把 $x=0$ 代入, 则分子分母都为 0. 尽管当 x 趋于 0 时, $\sin(x)$ 和 x 很接近, 但此时这一点对我们没有帮助, 因为我们正在考虑这两项的差, 所以我们使用洛必达法则. 首先分别对 $x-\sin(x)$ 和 x^3 求导：

$$\lim_{x\to 0}\frac{x-\sin(x)}{x^3}\overset{\text{l'H}}{=}\lim_{x\to 0}\frac{1-\cos(x)}{3x^2}.$$

在 7.1.2 节中, 我们实际上曾经演示过怎样求解等式右边的极限 (但原极限在分母中没有 3). 我们的原始方法是分子分母同时乘以 $1+\cos(x)$. 但是, 现在有一个更简单的方法：请注意, 当用 0 替代 x 时, 可发现该极限为 0/0 型 (因为 $\cos(0)=1$), 所以我们可以再次使用洛必达法则! 从而得到

$$\lim_{x\to 0}\frac{x-\sin(x)}{x^3}\overset{\text{l'H}}{=}\lim_{x\to 0}\frac{1-\cos(x)}{3x^2}\overset{\text{l'H}}{=}\lim_{x\to 0}\frac{\sin(x)}{6x}.$$

实际上, 可以多次使用洛必达法则计算最终的极限, 但对于这道题, 更好的写法是

$$\lim_{x\to 0}\frac{\sin(x)}{6x}=\frac{1}{6}\lim_{x\to 0}\frac{\sin(x)}{x}=\frac{1}{6}\times 1=\frac{1}{6}.$$

(在上述计算中, 我们直接使用了 7.1.5 节中的三角函数公式.) 总而言之, 我们得到了该极限的结果：

$$\lim_{x\to 0}\frac{x-\sin(x)}{x^3}=\frac{1}{6}.$$

介绍下一种形式之前, 让我们再重新观察一下这种形式. 回顾 6.5 节, 可以看到我们是用极限来定义导数的. 例如, 计算极限

$$\lim_{h\to 0}\frac{\sqrt[5]{32+h}-2}{h},$$

要使用一点技巧. 设 $f(x)=\sqrt[5]{x}$, 写出 $f'(x)$ 的表达式, 将其写成极限的形式, 最后把 $x=32$ 代入导函数的表达式. (检查一下细节.) 但是, 使用洛必达法则, 所有这些技巧就变得不必要了. 例如, 因为上述极限是 0/0 型, 所以可以同时对分子分母关于 h 求导来计算极限. 首先把 $\sqrt[5]{32+h}$ 改写为 $(32+h)^{1/5}$ 的形式, 这时有

$$\lim_{h\to 0}\frac{\sqrt[5]{32+h}-2}{h}=\lim_{h\to 0}\frac{(32+h)^{1/5}-2}{h}\overset{\text{l'H}}{=}\lim_{h\to 0}\frac{\frac{1}{5}(32+h)^{-4/5}}{1}=\frac{1}{5}\times(32)^{-4/5},$$



计算结果为 1/80, 这同我们之前计算的结果是相符的. 现在就请回到 6.5 节, 使用洛必达法则重新计算一下其中的例题.

14.1.2　类型 A: $\pm\infty/\pm\infty$



洛必达法则对于 $\lim_{x\to a} f(x)=\infty$ 和 $\lim_{x\to a} g(x)=\infty$ 的情况也很适用. 也就是说, 当把 $x=a$ 代入原函数时, 分子分母都趋于无穷大, 所以我们所求解的是 ∞/∞ 型. 例如, 求极限

$$\lim_{x\to\infty}\frac{3x^2+7x}{2x^2-5},$$

你能注意到, 当 $x\to\infty$ 时, 分子分母同时趋于 ∞, 所以可以使用洛必达法则:

$$\lim_{x\to\infty}\frac{3x^2+7x}{2x^2-5}\overset{\text{l'H}}{=}\lim_{x\to\infty}\frac{6x+7}{4x}=\lim_{x\to\infty}\left(\frac{6}{4}+\frac{7}{4x}\right).$$

当 $x\to\infty$ 时, $7/4x$ 趋于 0, 所以极限结果为 6/4, 也就是 3/2. 当然, 你也可以使用 4.3 节的方法计算极限. 你会发现, 无论用什么方法都会得到同样的结果 —— 3/2.



这儿还有另一个例子, 求

$$\lim_{x\to 0^+}\frac{\csc(x)}{1-\ln(x)}.$$

注意: 当 $x\to 0^+$ 时, 分子分母都趋于无穷大. 为什么呢? 因为当 $x\to 0$ 时, $\sin(x)$ 趋于 0, 所以 $\csc(x)$ 趋于无穷大; 同样, 当 $x\to 0^+$ 时, $\ln(x)\to-\infty$, 所以 $1-\ln(x)\to\infty$. 现在我们可以使用洛必达法则:

$$\lim_{x\to 0^+}\frac{\csc(x)}{1-\ln(x)}\overset{\text{l'H}}{=}\lim_{x\to 0^+}\frac{-\csc(x)\cot(x)}{-1/x}=\lim_{x\to 0^+}x\csc(x)\cot(x).$$

为了计算极限, 我们可以把它改写为

$$\lim_{x\to 0^+}\frac{x}{\sin(x)}\frac{1}{\tan(x)}.$$

这样有

$$\lim_{x\to 0^+}\frac{x}{\sin(x)}=\frac{1}{\lim\limits_{x\to 0^+}\dfrac{\sin(x)}{x}}=\frac{1}{1}=1,$$

但对于另一个因式有

$$\lim_{x\to 0^+}\frac{1}{\tan(x)}=\infty,$$

因为当 $x\to 0^+$ 时, $\tan(x)\to 0^+$, 所以我们证明了

$$\lim_{x\to 0^+}\frac{\csc(x)}{1-\ln(x)}=\infty.$$



当 $x\to\infty$ 时, 如我们前面看过的那样, 该法则也适用. 这里有另一个例子:

$$\lim_{x\to\infty}\frac{x}{\mathrm{e}^x}\overset{\text{l'H}}{=}\lim_{x\to\infty}\frac{1}{\mathrm{e}^x}=0.$$

该极限为 0, 因为当 $x \to \infty$ 时, $\mathrm{e}^x \to \infty$. 使用洛必达法则的前提条件是: 当 $x \to \infty$ 时, x 和 e^x 都趋于无穷大. 注意：分母 e^x 在求导的过程中是不变的, 但分子 x 的导数却为 1. 当你看到下述例子的时候, 可能会更清楚明了.

$$\lim_{x\to\infty} \frac{x^3}{\mathrm{e}^x}.$$

我们使用了三次洛必达法则, 发现每一次都是不定式 ∞/∞ 型：

$$\lim_{x\to\infty} \frac{x^3}{\mathrm{e}^x} \overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{3x^2}{\mathrm{e}^x} \overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{6x}{\mathrm{e}^x} \overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{6}{\mathrm{e}^x} = 0.$$

当然, 同样的方法可以应用到 x 的任何次幂; 但你不得不多次使用该法则, 每次都要求导至导数为 1 为止, 然而对于 e^x 无论求导多少次都不变. 其实在 9.4.4 节中, 我们已经仔细讨论并证明过指数函数增长得很快.

现在, 我有个非常善意的提示：请记住, 只有不定式才能用洛必达法则! 对于分式仅有的不定式是 0/0 或 $\pm\infty/\pm\infty$ 这两种形式. 例如, 如果你想对极限

$$\lim_{x\to 0} \frac{x^2}{\cos(x)}$$

使用洛必达法则, 将会陷入一团混乱. 让我们看看使用洛必达法则后的情况：

$$\lim_{x\to 0} \frac{x^2}{\cos(x)} \overset{\text{l'H?}}{=} \lim_{x\to 0} \frac{2x}{-\sin(x)} = -2 \lim_{x\to 0} \frac{x}{\sin(x)} = -2.$$

很显然, 这是错误的, 因为当 x 趋于 0 时, x^2 和 $\cos(x)$ 都为正. 事实上, 正确的解答过程是

$$\lim_{x\to 0} \frac{x^2}{\cos(x)} = \frac{0^2}{\cos(0)} = \frac{0}{1} = 0.$$

洛必达法则不能用来解答这道题, 原因是它是 0/1 型, 不是不定式. 所以, 一定要注意!

14.1.3 类型 B1: $(\infty - \infty)$

本章开始有这样一个极限表达式：

$$\lim_{x\to 0} \left(\frac{1}{\sin(x)} - \frac{1}{x}\right).$$

当 $x \to 0^+$ 时, $1/\sin(x)$ 和 $1/x$ 都趋于 ∞; 当 $x \to 0^-$ 时, 这两项又同时趋于 $-\infty$. 无论哪种情况, 这都是两个非常大的数 (正无穷大或负无穷大) 的差, 所以该不定式可以被表示为 $\pm(\infty - \infty)$.

幸运的是, 可以很容易地把这种形式转化为类型 A. 我们所需要做的仅仅是通分：

$$\lim_{x\to 0} \left(\frac{1}{\sin(x)} - \frac{1}{x}\right) = \lim_{x\to 0} \frac{x - \sin(x)}{x\sin(x)}.$$

现在, 把 x =0 代入, 可以看出这是 0/0 型不定式, 所以可以应用洛必达法则：

$$\lim_{x\to 0}\left(\frac{1}{\sin(x)}-\frac{1}{x}\right)=\lim_{x\to 0}\frac{x-\sin(x)}{x\sin(x)}\overset{\text{l'H}}{=}\lim_{x\to 0}\frac{1-\cos(x)}{\sin(x)+x\cos(x)}.$$

注意, 我们使用乘法法则对分母进行求导. 无论怎样, 再次出现了 0/0 型, 因为当把 $x=0$ 代入时, 可以发现分子分母同时变为 0. 所以再次使用洛必达法则 (并且再次使用乘法法则):

$$\lim_{x\to 0}\frac{1-\cos(x)}{\sin(x)+x\cos(x)}\overset{\text{l'H}}{=}\lim_{x\to 0}\frac{\sin(x)}{\cos(x)+\cos(x)-x\sin(x)}.$$

做到这儿, 就不要再使用洛必达法则了! 因为在本阶段, 当把 $x=0$ 代入时, 可发现分子为 0 但分母为 2, 所以极限结果为 0. 把刚才的计算综合到一起, 可得

$$\lim_{x\to 0}\left(\frac{1}{\sin(x)}-\frac{1}{x}\right)=0.$$

有时通分也不能解决问题. 比如你可能会遇到一道没有分母的题目, 所以不得不自己创造一个分母. 例如, 求极限

$$\lim_{x\to\infty}(\sqrt{x+\ln(x)}-\sqrt{x}),$$

首先注意, 当 $x\to\infty$ 时, $\sqrt{x+\ln(x)}$ 和 $\sqrt{x}$ 都趋于 ∞, 所以这道题属于 $\infty-\infty$ 不定式. 但该题目没有分母, 所以我们同时乘以除以一个共轭表达式:

$$\lim_{x\to\infty}(\sqrt{x+\ln(x)}-\sqrt{x})=\lim_{x\to\infty}(\sqrt{x+\ln(x)}-\sqrt{x})\times\frac{\sqrt{x+\ln(x)}+\sqrt{x}}{\sqrt{x+\ln(x)}+\sqrt{x}}.$$

通过使用平方差公式 $(a-b)(a+b)=a^2+b^2$, 有

$$\lim_{x\to\infty}\frac{x+\ln(x)-x}{\sqrt{x+\ln(x)}+\sqrt{x}}=\lim_{x\to\infty}\frac{\ln(x)}{\sqrt{x+\ln(x)}+\sqrt{x}}.$$

现在, 题目转化为类型**A**的 ∞/∞ 不定式, 所以通过对分子分母同时求导 (分母要使用链式求导法则) 可得

$$\lim_{x\to\infty}\frac{\ln(x)}{\sqrt{x+\ln(x)}+\sqrt{x}}\overset{\text{l'H}}{=}\lim_{x\to\infty}\frac{1/x}{\dfrac{1+1/x}{2\sqrt{x+\ln(x)}}+\dfrac{1}{2\sqrt{x}}}.$$

如果分子分母同时乘以 x, 可得

$$\lim_{x\to\infty}\frac{1}{\dfrac{x+1}{2\sqrt{x+\ln(x)}}+\dfrac{\sqrt{x}}{2}}.$$

我们差不多要得到答案了, 但仍需看看 $x\to\infty$ 时分母中第一个分式的值:

$$\lim_{x\to\infty}\frac{x+1}{2\sqrt{x+\ln(x)}}.$$

注意, 这也是 ∞/∞ 不定式, 所以需要再次使用洛必达法则:

$$\lim_{x\to\infty}\frac{x+1}{2\sqrt{x+\ln(x)}} \stackrel{\text{l'H}}{=} \lim_{x\to\infty}\frac{1}{\dfrac{2(1+1/x)}{2\sqrt{x+\ln(x)}}} = \lim_{x\to\infty}\frac{\sqrt{x+\ln(x)}}{1+1/x}.$$

当 $x\to\infty$ 时, 分母 $1+1/x$ 趋于 1, 但是分子 $\sqrt{x+\ln(x)}$ 趋于 ∞. 也就是说

$$\lim_{x\to\infty}\frac{x+1}{2\sqrt{x+\ln(x)}} = \infty.$$

回到原始的问题, 我们已经发现

$$\lim_{x\to\infty}(\sqrt{x+\ln(x)}-\sqrt{x}) = \lim_{x\to\infty}\frac{1}{\dfrac{x+1}{2\sqrt{x+\ln(x)}}+\dfrac{\sqrt{x}}{2}}.$$

当 $x\to\infty$ 时, 分母中的两个分式都趋于 ∞, 所以极限为 0.

不幸的是, 对于类型 **B1** 的极限, 洛必达法则并不是一直能解决问题. 事实上, 仅在你能把原始表达式转化为两式之比时它才有效.

14.1.4 类型B2: $(\mathbf{0}\times\pm\infty)$



下面这个出现在本章开头的极限, 其实我们在 9.4.6 节已经见过了:

$$\lim_{x\to 0^+} x\ln(x).$$

因为当 $x\leqslant 0$ 时, $\ln(x)$ 没有意义, 所以只需求当 $x\to 0^+$ 的极限. 可以看出, 当 $x\to 0^+$ 时, $x\to 0$ 然而 $\ln(x)\to -\infty$, 所以该极限为 $0\times(-\infty)$ 型不定式. 让我们通过处理分母把该极限转化为类型 A. 基本思想是把 x 转化为 $1/x$ 从而把 x 移到分母:

$$\lim_{x\to 0^+} x\ln(x) = \lim_{x\to 0^+}\frac{\ln(x)}{1/x}.$$

现在为 $-\infty/\infty$ 型, 所以可以使用洛必达法则:

$$\lim_{x\to 0^+} x\ln(x) = \lim_{x\to 0^+}\frac{\ln(x)}{1/x} \stackrel{\text{l'H}}{=} \lim_{x\to 0^+}\frac{1/x}{-1/x^2}.$$

最右边的极限可以化简为 $-x$, 最后的极限为

$$\lim_{x\to 0^+}(-x) = 0.$$

我们已经解决了问题, 但且回过头来再看看这道题: 为什么我把 x 移到分母而不是移动 $\ln(x)$ 呢? 如果移动 $\ln(x)$, 则为

$$\lim_{x\to 0^+} x\ln(x) = \lim_{x\to 0^+}\frac{x}{1/\ln(x)}.$$

现在, 你需要对 $1/\ln(x)$ 求导, 这个导数相对难求一点. 但如果你尝试一下, 可得

$$\lim_{x\to 0^+} x\ln(x) = \lim_{x\to 0^+}\frac{x}{1/\ln(x)} \stackrel{\text{l'H}}{=} \lim_{x\to 0^+}\frac{1}{(1/x)(-1/(\ln(x))^2)} = \lim_{x\to 0^+} -x(\ln(x))^2.$$

这比我们最原始的极限还复杂! 所以当把某一项移到分母时一定要注意. 从上述例子可以看出, 移动对数项是个很糟糕的思路, 所以要避免这样做.

这里还有一个例子：

$$\lim_{x\to\pi/2}\left(x-\frac{\pi}{2}\right)\tan(x).$$

当你把 $x=\pi/2$ 代入原函数时, 会发现第一项 $(x-\pi/2)$ 为 0, 而 $\tan(x)$ 这项要么为 ∞(当 $x\to(\pi/2)^-$), 要么为 $-\infty$(当 $x\to(\pi/2)^+$). 通过函数图像可以更加肯定我们的结论. 无论哪种情况, 都可以把 $\tan x$ 移到分母, 从而转化为 $1/\tan x$ 或 $\cot x$. 也就是：

$$\lim_{x\to\pi/2}\left(x-\frac{\pi}{2}\right)\tan(x)=\lim_{x\to\pi/2}\frac{x-\pi/2}{\cot(x)}.$$

这比把 $(x-\pi/2)$ 移到分母的计算量要小得多, 实际上, 把 $(x-\pi/2)$ 移到分母答案都算不出. 无论如何, 上述的极限形式都是 0/0 型, 所以可以使用洛必达法则：

$$\lim_{x\to\pi/2}\left(x-\frac{\pi}{2}\right)\tan(x)=\lim_{x\to\pi/2}\frac{x-\pi/2}{\cot(x)}\overset{\text{l'H}}{=}\lim_{x\to\pi/2}\frac{1}{(-\csc^2(x))}.$$

因为 $\sin(\pi/2)=1$, 所以可以得出 $\csc(\pi/2)=1$. 上述极限的结果为 -1.

14.1.5　类型C: $(1^{\pm\infty}, 0^0$或$\infty^0)$

最后, 我们研究最复杂的一种情况, 比如：

$$\lim_{x\to0^+}x^{\sin(x)},$$

这种形式的底和指数部分都带有变量 (在该例中为 x). 如果设 $x=0$, 那么我们得到 0^0, 这是不定式的另一种形式. 为求得该极限, 要使用类似于对数函数求导法则的一种方法 (参考 9.5 节). 基本思想是首先对 $x^{\sin(x)}$ 取对数, 接下来再求当 $x\to0^+$ 时的极限：

$$\lim_{x\to0^+}\ln(x^{\sin(x)}).$$

根据对数法则 (参见 9.1.4 节), 指数 $\sin(x)$ 可以移到对数的前面：

$$\lim_{x\to0^+}\ln(x^{\sin(x)})=\lim_{x\to0^+}\sin(x)\ln(x).$$

当 $x\to0^+$, 可得 $\sin(x)\to0$, $\ln(x)\to-\infty$, 所以该题属于类型 **B2**. 如果把 $\sin(x)$ 移到分母则为 $1/\sin(x)$, 也就是 $\csc(x)$, 这时该题又转化为类型 **A**, 这样就可以使用洛必达法则求解：

$$\lim_{x\to0^+}\sin(x)\ln(x)=\lim_{x\to0^+}\frac{\ln(x)}{\csc(x)}\overset{\text{l'H}}{=}\lim_{x\to0^+}\frac{1/x}{-\csc(x)\cot(x)}.$$

化简为

$$\lim_{x\to0^+}-\frac{\sin(x)}{x}\times\tan(x)=-1\times0=0.$$

做完了吗? 还没有. 我们现在知道：

$$\lim_{x\to 0^+} \ln(x^{\sin(x)}) = 0;$$

现在对两端同时求指数, 可得

$$\lim_{x\to 0^+} x^{\sin(x)} = \mathrm{e}^0 = 1.$$

(这种求指数的方法很有效果, 因为 e^x 是关于 x 的连续函数.)

我们对刚才的计算总结一下. 先不求原始函数的极限, 而是先对该函数取对数, 再用处理类型 **B2** 的方法计算极限. 最后, 再对刚才计算的结果求指数.

事实上, 有时我们需要用类型 **A** 而不是类型 **B2** 解决该问题. 例如, 求解

$$\lim_{x\to 0}(1+3\tan(x))^{1/x}.$$

这是本章开始部分的例题, 首先请注意我们正在求 $1^{\pm\infty}$ 类型的极限. 对其取对数有

$$\lim_{x\to 0}\ln\left((1+3\tan(x))^{1/x}\right) = \lim_{x\to 0}\frac{1}{x}\ln(1+3\tan(x)) = \lim_{x\to 0}\frac{\ln(1+3\tan(x))}{x}.$$

现在转化为 0/0 型了, 所以可以用类型 **A** 的方法解决问题了. 根据链式求导法则, 有

$$\lim_{x\to 0}\frac{\ln(1+3\tan(x))}{x} \overset{\text{l'H}}{=} \lim_{x\to 0}\frac{\dfrac{3\sec^2(x)}{1+3\tan(x)}}{1} = \frac{3(1)^2}{1+3(0)} = 3.$$

这样有

$$\lim_{x\to 0}\ln\left((1+3\tan(x))^{1/x}\right) = 3.$$

再对该等式两边同时求指数, 可得

$$\lim_{x\to 0}(1+3\tan(x))^{1/x} = \mathrm{e}^3.$$

这里还有另一个不定式 ∞^0 的例子:

$$\lim_{x\to\infty} x^{-1/x},$$

因为当 $x\to\infty$ 时, $-1/x\to 0$. 我们可以用同样的方法解决问题, 先对其取对数, 接下来再用洛必达法则求解 (用类型 **A** 的方法).

$$\lim_{x\to\infty}\ln(x^{-1/x}) = \lim_{x\to\infty}\frac{\ln(x)}{-x} \overset{\text{l'H}}{=} \lim_{x\to\infty}\frac{1/x}{-1} = 0.$$

再两边求指数, 可得

$$\lim_{x\to\infty} x^{-1/x} = \mathrm{e}^0 = 1.$$

我们需要知道的是关于指数类型的不定式不仅仅是 $1^{\pm\infty}$, 0^0 和 ∞^0. 你可以看出, 任何指数型函数都可以使用取对数的方法把问题转化为乘积或商的形式, 这时再求新的极限 L. 实际的极限结果将会是 e^L. 唯一的例外情况是 $L=\infty$, 这时将 e^∞ 替换为 ∞; 当 $L=-\infty$ 时, 就将 $\mathrm{e}^{-\infty}$ 看作 0. 这符合 9.4.4 节中的极限:

$$\lim_{x\to\infty}\mathrm{e}^x = \infty \quad 和 \quad \lim_{x\to-\infty}\mathrm{e}^x = 0$$

14.1.6　洛必达法则类型的总结

下面是我们已经讲解的所有技巧.

- **类型 A**　如果极限是分式的形式, 例如

$$\lim_{x\to a}\frac{f(x)}{g(x)},$$

要检查该形式是否为不定式. 该分式一定为 $0/0$ 或 $\pm\infty/\pm\infty$, 使用洛必达法则

$$\boxed{\lim_{x\to a}\frac{f(x)}{g(x)}\overset{\text{l'H}}{=}\lim_{x\to a}\frac{f'(x)}{g'(x)}.}$$

在求导的过程中, 请不要使用商的求导法则! 现在, 为求解这个新的极限, 可能需要再次使用洛必达法则.

- **类型 B1**　如果是求差的极限, 例如

$$\lim_{x\to a}(f(x)-g(x)),$$

该形式为 $\pm(\infty-\infty)$. 求解该极限的方法是通分或同时乘以除以一个共轭表达式从而转化为类型 **A**.

- **类型 B2**　如果极限是乘积的形式, 例如

$$\lim_{x\to a}f(x)g(x),$$

该形式为 $0\times\pm\infty$, 选择两个因式中较简单的那个取倒数把它移到分母 (尽量不要选用对数做分母, 把它留在分子). 这样就转化为

$$\lim_{x\to a}f(x)g(x)=\lim_{x\to a}\frac{g(x)}{1/f(x)}.$$

这是典型的类型 **A**.

- **类型 C**　如果极限为指数的形式, 并且该指数的底和指数部分都含变量, 例如

$$\lim_{x\to a}f(x)^{g(x)},$$

首先, 我们取其对数：

$$\lim_{x\to a}\ln(f(x)^{g(x)})=\lim_{x\to a}g(x)\ln(f(x)).$$

这样转化为类型 **B2** 或 **A** (或者转化后的结果不是不定式, 这时不得不想其他的技巧). 一旦你已经求解出来了, 这时, 就有

$$\lim_{x\to a}\ln(f(x)^{g(x)})=L,$$

然后再两边同时取指数, 可得

$$\lim_{x\to a}f(x)^{g(x)}=\mathrm{e}^L.$$

现在, 你需要做的就是多做习题来练习如何使用洛必达法则.

14.2 关于极限的总结

现在需要巩固一下我们所学的关于极限的知识. 下面简要总结了我们学过的关于计算极限的所有方法. 这些方法可应用于 $\lim_{x\to a} F(x)$ 形式的极限, F 是一个至少在 a 点附近连续的函数, 但在 a 点可能不连续. 当然, a 也可能是 ∞ 或 $-\infty$. 这样, 我们有如下的总结.

- **首先尝试使用替换法**. 这样你可能就会求得极限结果.
- 如果替换导致出现 b/∞ 或 $b/(-\infty)$ 的形式, b 是个限定的数, 那么该极限的结果为 0.
- 如果替换之后的形式为 $b/0$, 但 b 不为 0, 这时说明该函数有垂直渐近线, 即左极限和右极限为 ∞ 或 $-\infty$, 那么双侧极限或者不存在 (如果左右极限不相等) 或者为 ∞ 或 $-\infty$. 使用在 $x=a$ 点附近的符号表格去查找左极限和右极限 (请参见 4.1 节).
- 如果不是上述任何一种情况, 那么该极限就为 0/0 形式. 首先看它是否为导数定义的形式. 如果你可以把它改写为某种特定函数关于特定的数 x 的导数形式 $\lim_{h\to 0}\dfrac{f(x+h)-f(x)}{h}$, 这时该极限为 $f'(x)$. 我们在 14.1.1 节中见过该形式, 其实这种类型的极限也可以用洛必达法则来解决. (参见 6.5 节.)
- 如果极限有**根号**, 那么可以考虑分母有理化或分子有理化的方法. (参见 4.2 节.)
- 如果有**绝对值**, 那么要考虑把绝对值符号去掉, 即把该函数转化为分段函数的形式.

  $$|A| = \begin{cases} A & \text{如果 } A \geqslant 0, \\ -A & \text{如果 } A < 0. \end{cases}$$

  记得把上式中所有的五个 A 都替换为你正在处理的问题中需要去掉绝对值的具体表达式! (参见 4.6 节.)
- 另外可以利用不同函数的特性帮助你解决问题. 请记住在极限中, “无穷小” 意味着 “趋于 0”; “无穷大” 意味着正无限大的数或负无限大的数. (参见 3.4.1 节.) 请注意: 如果你所要求的是当 $x\to\infty$ 时的极限, 这并不能说明该极限就为无穷大. 例如, $\sin(1/x)$ 就是当 $x\to\infty$ 时函数值却越来越小的例子, 因为 $1/x\to 0$. 当 $x\to 0$ 时同样值得注意, 因为函数结果可能会是非常大的. 以下是关于多项式、三角函数、指数函数和对数函数的总结.

  (1) **多项式和多项式型函数**

  - **一般方法**　尝试因式分解, 然后把公因式约掉. (参见 4.3 节.)
  - **大讨论　最大次数的项决定该极限的值**, 所以同时除以并乘以该项. (参见 4.3 节.)

(2) **三角函数和反三角函数**

- **一般方法**　记住所有三角函数和反三角函数的图像, 以及它们在一些特殊点处的函数值. 这些知识点在第 2 章和第 10 章中有详细的讨论.
- **小讨论**　当 A 是个很小的数时, $\sin(A)$ 和 A 的数值非常接近, 所以可以乘以 A 并除以 A. 对 $\tan(A)$ 可以用同样的方法, 但 $\cos(A)$ 不可以, 因为当 A 趋于 0 时, $\cos(A)$ 趋于 1. 当仅涉及乘积或商的时候该方法很实用. 但当有三角函数的加减形式出现时, 该方法可能就不管用了. (参见 7.1.2 节.)
- **大讨论**　对于正弦或余弦函数, 我们可以利用它们的特性 $|\sin(\text{任意数})| \leqslant 1$ 和 $|\cos(\text{任意数})| \leqslant 1$. 这个特性可以同三明治定理一起使用. (参见 7.1.3 节.) 这里也有一些其他有用的特性
$$\lim_{x\to\infty} \tan^{-1}(x) = \frac{\pi}{2} \quad 和 \quad \lim_{x\to-\infty} \tan^{-1}(x) = -\frac{\pi}{2}.$$
(非正式地, 你可以这样来记 $\tan^{-1}(\infty) = \pi/2$ 和 $\tan^{-1}(-\infty) = -\pi/2$, 但要确保你真正理解了上述这两个公式的含义.)

(3) **指数函数**

- **一般方法**　要记住 $y = \mathrm{e}^x$ 的图像, 也要知道下列两个极限:
$$\lim_{h\to 0}(1 + hx)^{1/h} = \mathrm{e}^x \quad 和 \quad \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n = \mathrm{e}^x.$$
(参见 9.4.1 节.)
- **小讨论**　因为 $\mathrm{e}^0 = 1$, 所以当极限的表达式中有该因式时, 完全可以用 1 替代它. 但当极限中有和或差的形式时, 问题就不是这么简单了. 这时, 你不得不考虑使用洛必达法则或用导数的定义去求解. (参见 9.4.2 节.)
- **大讨论**　记住以下这两个重要的极限:
$$\lim_{x\to\infty} \mathrm{e}^x = \infty \quad 和 \quad \lim_{x\to-\infty} \mathrm{e}^x = 0.$$
(仅仅为替换的目的, 可以把这两个极限考虑为 $\mathrm{e}^{\infty} = \infty$ 和 $\mathrm{e}^{-\infty} = 0$, 尽管这两个等式并不符合正式的写法.)当然也要记住, 当 $x \to \infty$ 时, 指数函数增长得很快. 也就是说 $\lim\limits_{x\to\infty} \dfrac{\text{多项式}}{\mathrm{e}^x} = 0$. 底 e 可以是任何大于 1 的数, 指数 x 可以是任何最高项系数为正数的多项式. (参见 9.4.4 节.)

(4) **对数函数**

- **一般方法**　需要知道 $y = \ln(x)$ 的图像以及对数的运算法则, 这些知识在 9.1.4 节中已经介绍过.

- **小讨论** 一个很重要的极限是

  $$\lim_{x \to 0^+} \ln(x) = -\infty$$

  (或者也可以这样记, $\ln(0) = -\infty$). 另外, 当 $x \to 0^+$ 时, 对数函数慢慢地趋于 $-\infty$, 即对于任何大于 0 的数 a, 无论 a 有多小 (参见 9.4.6 节) 都有

  $$\lim_{x \to 0^+} x^a \ln(x) = 0$$

- **大讨论** 我们有

  $$\lim_{x \to \infty} \ln(x) = \infty,$$

  也可以非正式地简写为 $\ln(\infty) = \infty$. 不管怎样说, **对数增长得很慢**, 比任何多项式都慢, 即对于任何次数为正的多项式 (参见 9.4.5 节) 都有

  $$\lim_{x \to \infty} \frac{\ln(x)}{\text{多项式}} = 0.$$

- **函数在 1 附近的情况** 我们有 $\ln(1) = 0$. 此时, 洛必达法则可能是很有用的, 对于这种极限或者也可以使用极限的定义来解决. (参见 9.4.3 节.)

- 如果上述方法都不能解决问题, 那么考虑使用洛必达法则 (参见 14.1.6 节的总结). 在使用洛必达法则以后, 总是会得到一个新的极限, 我们可以考虑使用上述的任何方法或再次使用洛必达法则.

上述所有的事实及方法仅仅是求极限的工具, 它们不可能解决所有的极限问题. 事实上, 在第 17 章中, 我们将看到一种完全不同的求解极限的方法. 如果你对现在的知识掌握得很扎实, 将会有助于你以后的学习. 知道用哪种方法去解决问题是一种艺术, 当然, 熟能生巧. 所以要多做练习!

第 15 章　积　　分

就微积分问题, 微分仅仅是其中的一半内容, 另一半是积分. 积分是个很强大的数学工具, 它可以帮助我们求不规则图形的面积, 不规则形状物体的体积, 以及变速运动物体的路程. 本章中, 我们会花一些时间去研究定义定积分所需要的理论. 在下一章, 将给出其定义并研究其应用. 所以让我们从预备知识入手：

- 求和符号以及伸缩求和法;
- 寻求位移和面积之间的关系;
- 用分割法去求面积.

15.1　求 和 符 号

考虑求和

$$\frac{1}{1}+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\frac{1}{36}.$$

这并不是随意数的求和, 这是有一定规律的求和. 和式的每一项都是平方数的倒数. 这里有个更便捷的方法表达这个求和：

$$\sum_{j=1}^{6}\frac{1}{j^2}.$$

请大声地读出："从 $j=1$ 到 $j=6$ 时 $1/j^2$ 的和." 现在我要介绍这个求和符号是怎样工作的. 思路是把 $j=1$, $j=2$, $j=3$, $j=4$, $j=5$ 以及 $j=6$ 代入 $1/j^2$, 一次代入一个, 最后把这六项都加到一起. 我们从 $j=1$ 开始, 以 $j=6$ 结束, $j=1$ 和 $j=6$ 分别在这个希腊字母 $\sum$ 的上方和下方. ($\sum$ 是希腊字母 sigma 的大写, 我们叫这个符号为 "sigma 求和".) 所以我们有

$$\sum_{j=1}^{6}\frac{1}{j^2}=\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\frac{1}{5^2}+\frac{1}{6^2}.$$



注意：我们实际上并没有计算出这个和的值! 所做的仅仅是缩写和式.

现在, 考虑下边这个使用求和符号的级数 (这是 "求和" 的另一个说法):

$$\sum_{j=1}^{1000}\frac{1}{j^2}.$$

这个求和与上个求和之间的唯一区别是, 这次累加到 1000 而不是 6 了. 所以

$$\sum_{j=1}^{1000}\frac{1}{j^2}=\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{999^2}+\frac{1}{1000^2}.$$

在这种情况下, 求和符号是非常实用的, 它避免了使用"$\cdots$". 这里还有另一个例子:

$$\sum_{j=5}^{30}\frac{1}{j^2}=\frac{1}{5^2}+\frac{1}{6^2}+\frac{1}{7^2}+\cdots+\frac{1}{29^2}+\frac{1}{30^2}.$$

这个求和是以 $j=5$ 开始, 而不是 $j=1$, 所以第一项是 $1/5^2$.

当你要考虑改变求和的起点和终点时, 求和符号是非常方便的. 例如, 考虑下边这个级数:

$$\sum_{j=1}^{n}\frac{1}{j^2}.$$

从 $j=1$ 开始, 以 $j=n$ 结束, 所以我们有

$$\sum_{j=1}^{n}\frac{1}{j^2}=\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{(n-2)^2}+\frac{1}{(n-1)^2}+\frac{1}{n^2}.$$

注意: 上边的等式倒数第二项对应 $j=n-1$, 倒数第三项对应 $j=n-2$; 我把后三项和前三项都写在了等式的右边, 而其他的项用"$\cdots$"符号在中间代替. 求和式

$$\sum_{j=1}^{n}\frac{1}{j^2}$$

看起来好像有两个变量 j 和 n, 但实际上只有一个变量 n. 把它展开, 可以容易地看出这一点

$$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{(n-2)^2}+\frac{1}{(n-1)^2}+\frac{1}{n^2}.$$

在这个展开式中没有 j! j 只是虚拟变量, 是个临时的替代者, 我们把它叫作**求和指标**, 它遍历从整数 1 到整数 n 之间的所有整数. 所以我们可以换用另一个字母, 而对表达式毫无影响. 例如, 下述求和是完全一样的:

$$\sum_{j=1}^{6}\frac{1}{j^2}=\sum_{k=1}^{6}\frac{1}{k^2}=\sum_{a=1}^{6}\frac{1}{a^2}=\sum_{\alpha=1}^{6}\frac{1}{\alpha^2}.$$

顺便说一下, 这已经不是我们第一次使用像 j 一样的虚拟变量了: 极限也使用这样的变量, 所以这里没有什么新鲜的. (参见 3.1 节的结尾部分.)

让我们看更多的例子.

$$\sum_{m=1}^{200}5$$

这个求和的结果是多少呢? 如果你说结果是 5, 那么你就掉入陷阱了. 我们仔细研究研究. 当 $m=1$ 时, 该项为 5; 当 $m=2$ 时, 它所对应的项还是 5. 对于 $m=3$,

$m = 4$ 直到 $m = 200$ 都是同样的结果. 所以, 实际上：

$$\sum_{m=1}^{200} 5 = 5 + 5 + 5 + \cdots + 5 + 5 + 5,$$

这个求和中有 200 项, 最后的结果应该是 200×5 或 1000. 类似地, 考虑下边这个求和：

$$\sum_{q=100}^{1000} 1 = 1 + 1 + 1 + \cdots + 1 + 1 + 1.$$

这个求和中一共有多少个 1? 你可能被诱导地说有 $1000 - 100$ 个或 900 个, 但实际上你少说了 1 个. 答案是 901 个. 总的来说, 在 A 和 B 之间, 包括 A 和 B, 共有 $B - A + 1$ 个数.

现在给你提个问题, 怎样用求和符号写下边的表达式：

$$\sin(1) + \sin(3) + \sin(5) + \cdots + \sin(2997) + \sin(2999) + \sin(3001).$$

你可能会写成

$$\sum_{j=1}^{3001} \sin(j),$$

但这并不正确, 按照这个写法展开后应该是

$$\sin(1) + \sin(2) + \sin(3) + \cdots + \sin(2999) + \sin(3000) + \sin(3001).$$

原展开式中没有偶数部分. 我们怎样才能去掉偶数部分呢? 首先想像 j 是从 1, 2, 3 开始遍历自然数, 这时 $(2j-1)$ 恰好遍历所有奇数. 所以我们把它写为

$$\sum_{j=1}^{3001} \sin(2j - 1).$$

这次好多了, 但仍然有个问题. j 的终点是 3001, 此时 $(2j-1)$ 为 $2\times(3001)-1 = 6001$. 也就是说

$$\sum_{j=1}^{3001} \sin(2j - 1) = \sin(1) + \sin(3) + \sin(5) + \cdots + \sin(5997) + \sin(5999) + \sin(6001).$$

这里有太多项了! 怎样知道在哪里停下呢? $\sin(2j - 1)$ 的最后一项, 应为 $\sin(3001)$ 而不是 $\sin(6001)$, 所以依据 $2j - 1 = 3001$, 可得 $j = 1501$. 最后我们有

$$\sin(1) + \sin(3) + \sin(5) + \cdots + \sin(2997) + \sin(2999) + \sin(3001) = \sum_{j=1}^{1501} \sin(2j - 1).$$

上述解答是正确的. 每次做完一定要把 j 代入校验, 本题中, 我们代入 j =1, j =2, j =3 及后三项 $j = 1499$, $j = 1500$ 和 $j = 1501$. 可以发现, 等式左右两侧是一样的. 此外, 我们再研究一下当 j 为偶数时的情况,

$$\sum_{j=1}^{1501}\sin(2j)$$

展开后为

$$\sin(2)+\sin(4)+\sin(6)+\cdots+\sin(2998)+\sin(3000)+\sin(3002).$$

所以, 当要得到偶数时, 我们使用 $2j$ 而不是 $(2j-1)$. 当然如果你要得到 3 的倍数, 应该使用 $3j$. 可能性是永无止境的!

15.1.1 一个有用的求和

考虑求和

$$\sum_{j=1}^{100} j.$$

首先, 展开这个求和. 当 j =1 时, 得到 1; 当 j =2 时, 得到 2; 以此类推, 直到 $j=100$; 这时, 仅仅需要把这 100 个数加起来. 所以有

$$\sum_{j=1}^{100} j = 1+2+3+\cdots+98+99+100.$$

是的, 这就是前 100 个自然数的和. 现在, 我们考虑下边这个求和

$$\sum_{j=0}^{99}(j+1).$$

当 j =0 时, 得到 1; 当 j =1 时, 得到 2; 同样, 以此类推, 直到 j =99, 得到 100. 所以, 实际上

$$\sum_{j=0}^{99}(j+1) = 1+2+3+\cdots+98+99+100.$$

这个求和同刚刚那个是一样的! 我们所做的只是把求和指标 j 减小了 1. 现在考虑这个求和

$$\sum_{j=1}^{100}(101-j).$$

当 j =1 时, 得到 100; 当 j =2 时, 得到 99; 以此类推, 到 j =100, 最后一项为 1. 也就是 $101-j$ 这个数从 100 递减到 1, 所以

$$\sum_{j=1}^{100}(101-j) = 100+99+98+\cdots+3+2+1.$$

这依然与刚才的两个求和一样, 只是这次是倒过来写罢了. 其实对于同一个求和用求和符号来表达会有很多种方法.

实际上, 这个求和的最后一个表达形式很普通, 我们可以用它来求出实际的数值. 假设 S 为 $1+2+3+\cdots+99+100$, 这时可以看到

$$S=\sum_{j=1}^{100} j \quad \text{和} \quad S=\sum_{j=1}^{100}(101-j).$$

如果把这两个求和加到一起, 就有

$$2S=\sum_{j=1}^{100} j+\sum_{j=1}^{100}(101-j).$$

在第一个求和中, 我们从 1 递增到 100; 而在第二个求和中, 我们从 100 递减到 1. 也就是说, 能够以任何顺序求得这个和而得到同样的结果. 所以, 可把这两个数合并到一起写为

$$2S=\sum_{j=1}^{100}(j+(101-j)).$$

因为 $j+(101-j)=101$, 这样, 结果会为

$$2S=\sum_{j=1}^{100} 101.$$

我们一共有 100 个 101, 所以有 $2S=101\times 100=10\,100$, 也就是 $S=10\,100/2=5050$. 这样就证明了从 1 加到 100 的和为 5050. 无论你信不信, 伟大的数学家高斯在 10 岁的时候就是用同样的方法解决该问题的!

15.1.2　伸缩求和法

检查求和

$$\sum_{j=1}^{5}(j^2-(j-1)^2).$$

它完全扩展后为

$$(1^2-0^2)+(2^2-1^2)+(3^2-2^2)+(4^2-3^2)+(5^2-4^2).$$

在这个求和中可以消掉很多相同的项. 实际上, 如果你仔细观察, 就会发现除了 5^2-0^2 之外的每一项都会被消掉, 所以求和的结果就是 $5^2=25$. 即使你有更多的项, 也是如此. 例如



$$\sum_{j=1}^{200}(j^2-(j-1)^2)$$

扩展后为

$$(1^2-0^2)+(2^2-1^2)+(3^2-2^2)+\cdots+(198^2-197^2)+(199^2-198^2)+(200^2-199^2).$$

再一次, 除了 200^2-0^2 之外的每一项都被消掉了, 所以这个和为 40 000. 等一下, 好像 3^2 和 -197^2 并没有被消掉! 其实它们隐藏在 $\cdots$ 里面了, 所以这个消元法的确奏效了.

这种类型的级数叫**伸缩级数**. 你可以把它合并成更简单的形式, 就像套缩那些老式的小望远镜一样. 总的来说, 我们有

$$\boxed{\sum_{j=a}^{b}(f(j)-f(j-1))=f(b)-f(a-1).}$$

例如,

$$\sum_{j=10}^{100}\left(\mathrm{e}^{\cos(j)}-\mathrm{e}^{\cos(j-1)}\right)=\mathrm{e}^{\cos(100)}-\mathrm{e}^{\cos(10-1)}$$

可以简单地写为 $\mathrm{e}^{\cos(100)}-\mathrm{e}^{\cos(9)}$. 我们只需取 $\mathrm{e}^{\cos(j)}$ 这项, 用最后的数 (100) 去替代 j, 用所得到的结果减去 $\mathrm{e}^{\cos(j-1)}$ 这项, 其中的 j 用数 (10) 去代替. 你应该把这个求和展开, 然后看看消元法是否帮助你得到了正确的答案.

还有另一个例子. 求

$$\sum_{j=1}^{n}(j^2-(j-1)^2),$$

注意, 这是个伸缩求和; 所以只需要取最后一项 $(j^2-(j-1)^2)$ 并用 n 去替代第一个 j, 以及用 1 去替代第二个 j, 可得

$$\sum_{j=1}^{n}(j^2-(j-1)^2)=n^2-(1-1)^2=n^2.$$

另一方面, $(j^2-(j-1)^2)$ 这项化简后为 $(j^2-(j^2-2j+1))$, 即 $2j-1$. 所以, 我们实际上证明了

$$\sum_{j=1}^{n}(2j-1)=n^2.$$

仔细考虑这个求和, 会发现左边仅仅是前 n 个奇数的和. 例如当 $n=5$ 时, 左边是 1+3+5+7+9, 这个和是 25. 这就是 5^2! 如果换个数取 $n=6$, 这时左边为 1+3+5+7+9+11, 这个和为 36, 正好是 6^2. 这再次证明了我们的结论是正确的. 这样已经证明了前 n 个奇数的和为 n^2.

我们甚至可以举更多的例子. 可以把这个求和分解为

$$\sum_{j=1}^{n}(2j)-\sum_{j=1}^{n}1=n^2.$$

如果你怀疑这个表达式, 请用前五项去校验一下. 不用常规的写法 1+3+5+7+9, 这次我们写为 $(2-1)+(4-1)+(6-1)+(8-1)+(10-1)$, 然后再重新安排一下得 $(2+4+6+8+10)-(1+1+1+1+1)$. 实际上, 可以从第一个括号中提出一个 2, 这样可得 $2\times(1+2+3+4+5)$. 根据上述的等式, 说明可以把常数 2 从第一个和中提出并得到

$$2\sum_{j=1}^{n} j - \sum_{j=1}^{n} 1 = n^2.$$

把第二个和移到等式的右边, 这时, 等式的两边同时除以 2, 可得

$$\sum_{j=1}^{n} j = \frac{1}{2}\left(n^2 + \sum_{j=1}^{n} 1\right).$$

最右边的求和是 n 个 1, 所以它实际上是 n. 等式右边为 $(n^2+n)/2$, 也可以被写为 $n(n+1)/2$. 这样我们证明了这个有用的公式

$$\sum_{j=1}^{n} j = \frac{n(n+1)}{2}.$$

当 $n = 100$ 时, 该公式为

$$\sum_{j=1}^{100} j = \frac{100(100+1)}{2} = 5050,$$

这同上一节的结论是一样的.

在刚才的例子中, 我们已经介绍了平方项, 现在来看看立方的情况：

$$\sum_{j=1}^{n} (j^3 - (j-1)^3) = n^3 - (1-1)^3 = n^3.$$

再一次, 由于这是个伸缩求和, 所以会很容易求出这个求和的值. 不管怎样, 你可以做一些代数运算, 会发现 $j^3 - (j-1)^3$ 化简后为 $3j^2 - 3j + 1$. 所以上述的求和为

$$\sum_{j=1}^{n} (3j^2 - 3j + 1) = n^3.$$

让我们把这个和分成三部分并把常数部分提出来：

$$3\sum_{j=1}^{n} j^2 - 3\sum_{j=1}^{n} j + \sum_{j=1}^{n} 1 = n^3.$$

现在, 把最后的两个和移到等式的右侧再除以 3, 可得

$$\sum_{j=1}^{n} j^2 = \frac{1}{3}\left(n^3 + 3\sum_{j=1}^{n} j - \sum_{j=1}^{n} 1\right).$$

上一个例子已经证明了等式右端的第一个和的结果为 $n(n+1)/2$; 第二个和为 n 个 1, 即为 n. 所以我们有

$$\sum_{j=1}^{n} j^2 = \frac{1}{3}\left(n^3 + \frac{3n(n+1)}{2} - n\right).$$

通过一些简单的代数运算可以得出等式右边的多项式可以化简为 $(2n^3 + 3n^2 + n)/6$, 因式分解后为 $n(n+1)(2n+1)/6$. 所以我们已经证明了

$$\sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6}.$$

现在, 我们就知道了怎样求前 n 个数的平方和. 例如

$$1^2+2^2+3^2+\cdots+99^2+100^2 = \frac{(100)\times(101)\times(201)}{6} = 338\,350.$$

即使是伟大的数学家高斯也要等到 11 岁才能解决这个问题!

15.2 位移和面积

介绍了求和符号之后, 我们来花一些时间研究下面这个问题：

> 如果你知道一辆汽车在某一时段内每一时刻的行驶速度, 那么它在这个时间段内的总位移是多少呢？

用符号来说明就是, 我们知道在 $[a, b]$ 时间段内每一时刻 t 的速度 $v(t)$, 想要求出它的总位移 $x(t)$. 我们已经知道反过来怎样计算：如果知道 $x(t)$, 那么 $v(t)$ 就是 $x'(t)$. 也就是说, 速度是位移对时间的导数. 为了解答这个问题, 首先让我们看一些简单的例子.

15.2.1 三个简单的例子

考虑三辆车沿着一条笔直的高速公路向前行驶. 因为车一直都是向前行驶, 所以可以用速率和路程来分别代替速度和位移 (对于这个情况, 这两个说法没有区别). 每一辆车都是在下午 3 点钟离开加油站, 下午 5 点钟结束旅行.

第一辆车匀速行驶, 在整个时间段内的平均速率为 50 英里/小时. 所以在 $[3, 5]$ 这个时间段内的速率为 $v(t) = 50$. 很容易就能计算出这辆车所走的路程, 我们仅仅需要使用这个公式：路程 = 平均速率 × 时间. 幸运的是, 对于这道题, 因为是匀速运动, 所以平均速率 v_{av} 和即时速率 v 是一样的, 都为 50. 所以有

$$\text{路程} = v \times t = 50 \times 2 = 100.$$

也就是说, 这辆车一共走了 100 英里. 现在让我们绘制速率 v 对时间 t 的图像, 如图 15-1 所示.

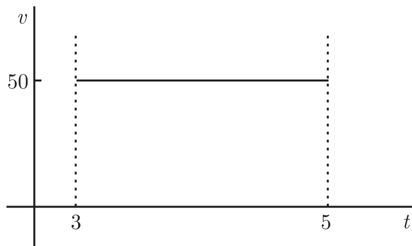


图 15-1

我们可以发现, 在 v =50 这条速度线以及时间轴 t =3 和 t =5 两条垂直线之间的图形是长方形. 长方形的高就是这辆车的速率 50 (英里/小时), 底就是它一共行驶的时间 2 (小时). 50×2 这个数值就是这个长方形的面积 (英里, 让我们暂时先不考虑这个单位). 所以在这个情况下这辆车所走的路程就是速度对时间的图像的面积.

接下来介绍第二辆车, 它在第一个小时的速度为 40 英里/小时; 从 4 点钟开始它以 60 英里/小时的速度行驶. 注意：我们忽略加速的那几秒钟, 那么它的图像如图 15-2 所示.

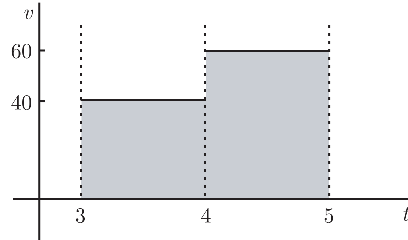


图　15-2

我已经把速率对时间的函数图像在 t =3 和 t =5 之间的部分用阴影表示了出来, 希望这就是路程. 让我们校验一下. 在第一个小时内, 它的速率为 40 英里/小时, 所以它所走的路程为 $40 \times 1 = 40$ 英里. 这恰恰就是该图像左长方形的面积, 高为 40(英里), 底为 1(小时). 对于第二小时, 可以用同样的方法进行分析. 它所走的路程为 60×1=60 英里 —— 这同右长方形的面积是一样的. 总面积没变, 还是 100.

重要的事实是, 我们根据该车的运动速度把它的运动时间分成几个时间段, 然后再求每个时间段的路程, 最后把这些时间段的路程加到一起. 类似于 $d = v_{\text{平均}} \times t$ 的公式并不适用于整段旅行, 除非你知道它的平均速度. 等一下, 可能你会说它的平均速度明显为 50 英里/小时, 所以解决这道题没有问题! 很好, 确实如此! 让我们看看第三个例子, 你是否还会有同感.

第三辆车在开始 15 分钟的速率为 20 英里/小时, 接下来一直到 4 点钟的速率为 40 英里/小时. 在 4 点钟的时候, 它提速到 60 英里/小时, 该速率保持了半个小时. 在最后的半个小时中, 它的速率降为 50 英里/小时. 我们再次忽略短暂的变速过程. 这样速率对时间的图像如图 15-3 所示.

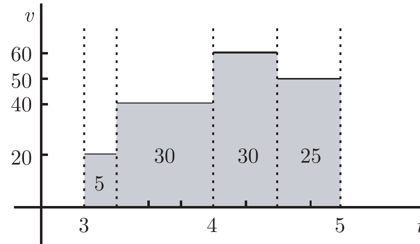


图　15-3

从这个图像中可以看出, 平均速率并不容易求出. 但我们仍可以通过把这 2 个小时的时间段分成 4 个小时间段去计算出它所走的路程. 通过图像可以看出, 共有 4 个长方形:

- 从 3 到 3.25(这是下午 3:15 的十进制表示法), 该车的速率为 20 英里/小时, 所以它所走过的路程是 20×0.25=5 英里. 这就是该图像的第一个长方形的面积, 因为它的高为 20(英里/小时), 底为 0.25(小时).
- 从 3.25 到 4, 它的速率为 40 英里/小时, 所以所走的路程为 40×0.75, 即 30 英里. 这也恰恰是第二个长方形的面积.
- 从 4 到 4.5(也就是下午 4:30), 这辆车的速率为 60 英里/小时, 所以路程为 60×0.5=30 英里, 这正是第三个长方形的面积.
- 最后从 4.5 到 5, 它的速率为 50 英里/小时, 所以在那段时间所走的路程为 50×0.5=25 英里, 这也是第四个长方形的面积.

像上图显示的那样, 在这四个时间段内, 这辆车分别行使了 5, 30, 30 和 25 英里, 所以它一共行驶了 5+30+30+25=90 英里. 最后, 我们也求出了第三辆车所走的路程! 这说明它的平均速率实际上是 90/2=45 英里/小时, 不是这四个时间段中的任意一个速率. (这并不违背中值定理, 因为上述函数是不可导的.)

15.2.2 一段更常规的旅行

我们再看一个描述这三辆车行驶过程的一般框架. 假设时间段为 $[a, b]$, 并且也假设这个时间段可以分成许多个更小的时间段, 从而保证汽车在每个小的时间段内是匀速行驶的. 我们不想固定时间段的数目, 所以假设共有 n 段. 我们也需要一些方法去描述每个时间段的开始和结束.

- 第一个时间段从时刻 a 开始, 以后来的某一时刻 t_1 结束. 因为 a 是比 t_1 更早的时刻, 所以可以说 $a < t_1$. 实际上, 如果设 $t_0 = a$, 那么对我们以后的解题会更有帮助, 所以有 $t_0 = a < t_1$.
- 第二个时间段从 t_1 时刻开始, 以后来的某时刻 t_2 结束, 这样, 我们有 $t_1 < t_2$.
- 第三个时间段从 t_2 到 t_3, 且 $t_2 < t_3$.
- 按照这个思路做下去, 所以到第 j 个时间段时, 我们是从 t_{j-1} 开始以 t_j 结束.
- 倒数第二个时间段从 t_{n-2} 到 t_{n-1}, 其中 $t_{n-2} < t_{n-1}$.
- 最后一个时间段从 t_{n-1} 到 t_n, t_n 同 b 时刻是一样的. 所以, 我们有 $t_{n-1} < t_n = b$.

综上所述, 我们可以说

$$a = t_0 < t_1 < t_2 < t_3 < \cdots < t_{n-2} < t_{n-1} < t_n = b.$$

我们已经把时间段 $[a, b]$ 分成了许多小时间段, 我们叫这样的小时间段为**分区**. 在

数轴上可表示为



在图像中间的这些小点表示我们并没有限定分区的个数.

除了考虑时间因素之外, 我们还需要考虑速度因素. 让我们假设在第一个小时间段 (t_0, t_1) 内, 汽车的行驶速度为 v_1. 这也就是说在 t_0 到 t_1 时间段内速度 v 对时间 t 的函数图像将是一条高度为 v_1 的线段. 对于第二时间段, 速度为 v_2, 所以对于该时间段的函数图像就得到一条高度为 v_2 的线段. 以此类推, 直到最后一个时间段 (t_{n-1}, t_n), 速度是 v_n. 所以图像大致如图 15-4 所示.

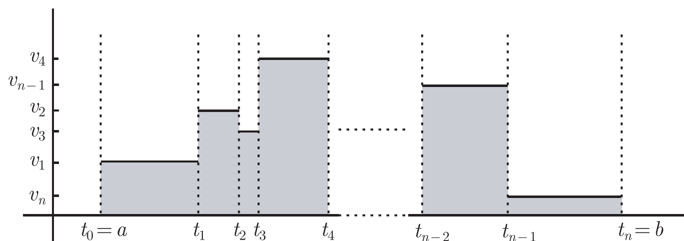


图　15-4

现在我们已经做好了计算总位移的准备. 在第一个小的时间段 (t_0, t_1) 内, 该车的速度为 v_1. 时间的长度为 $(t_1 - t_0)$, 所以在该时间段内所走过的位移为 $v_1 \times (t_1 - t_0)$. 让我们用同样的方法来计算第二时间段 (t_1, t_2) 的位移. 速度为 v_2, 该段时间的长度为 $(t_2 - t_1)$, 所以该时间段内的位移为 $v_2 \times (t_2 - t_1)$. 用同样的方法计算下去直到最后一个时间段 (t_{n-1}, t_n). 最后, 把所有的位移加到一起, 可得

$$\text{总位移} = v_1(t_1 - t_0) + v_2(t_2 - t_1) + \cdots + v_{n-1}(t_{n-1} - t_{n-2}) + v_n(t_n - t_{n-1}).$$



此时用 $\sum$ 来表达这个求和表达式将会非常恰当, 像我们在 15.1 节中使用过的那样. 校验看看, 使自己相信我们可以将上述公式写成如下形式：

$$\text{总位移} = \sum_{j=1}^{n} v_j(t_j - t_{j-1}).$$

当然, 这也是上述函数图像中阴影部分的面积.

来看看我们给出的这个框架是怎样适用于刚才的三个例子的. 对于每一个例子, 我们都有 $a = 3$ 和 $b = 5$.

- 对于第一辆车, 时间段为 $[3, 5]$, 所以设 $n = 1$, $t_0 = 3$ 和 $t_1 = 5$. 我们也知道它的速度为 $v_1 = 50$; 所以

$$\text{位移} = \sum_{j=1}^{n} v_j(t_j - t_{j-1}) = v_1(t_1 - t_0) = 50 \times (5 - 3) = 100.$$

- 第二辆车需要两个时间段; 设 n =2, $t_0 = 3$, $t_1 = 4$ 和 $t_2 = 5$, 所以我们的分区为 $3 < 4 < 5$. 在第一个时间段内, 速度 $v_1 = 40$, 在第二个时间段内速度 $v_2 = 60$. 所以

$$\text{位移} = \sum_{j=1}^{n} v_j(t_j - t_{j-1}) = v_1(t_1 - t_0) + v_2(t_2 - t_1)$$
$$= 40 \times (4 - 3) + 60 \times (5 - 4) = 100.$$

- 最后, 请你来分析第三辆车的运动的各个细节. 我们可以说 $n = 4$, 该分区为 $3 < 3.25 < 4 < 4.5 < 5$, 速度分别为 $v_1 = 20, v_2 = 40, v_3 = 60$ 和 $v_4 = 50$, 所以

$$\begin{aligned}\text{位移} &= \sum_{j=1}^{n} v_j(t_j - t_{j-1}) \\ &= v_1(t_1 - t_0) + v_2(t_2 - t_1) + v_3(t_3 - t_2) + v_4(t_4 - t_3) \\ &= 20 \times (3.25 - 3) + 40 \times (4 - 3.25) + 60 \times (4.5 - 4) + 50 \times (5 - 4.5) \\ &= 5 + 30 + 30 + 25 = 90.\end{aligned}$$

该计算同上一节的计算完全一样, 仅仅是其中的符号有些改变.

15.2.3 有向面积

如果我们的车向相反的方向行驶, 结果又会是怎样呢? 例如, 假设该车从下午 3 点到下午 4 点向正前方向行驶, 速度为 40 英里/小时, 然后以 30 英里/小时的速度向相反的方向行驶直到下午 6 点钟. 那么它的速度对时间的函数图像如图 15-5 所示.

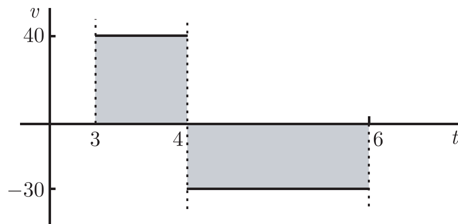


图 15-5

现在, 区分路程和位移真的很重要了. 在下午 3 点到 4 点之间, 路程和位移都为 40 英里. 从 4 点到 6 点, 该车公行驶了 $30 \times 2 = 60$ 英里, 所以从 3 点到 6 点的总路程为 40+60=100 英里. 另一方面, 因为该行驶过程的第二部分为反向行驶, 所以它的位移却为 $40 + (-60) = -20$ 英里. 这说明这辆车最终离出发点相反的方向 20 英里.

现在请看上边的函数图像. 左边长方形的面积为 40(英里), 这个不是问题; 但右边长方形的面积很有趣, 它的底长 2(小时), 如果你认为它的高为 30(英里/小时),

这时, 会足够确信它的面积为 60(英里). 把这两个面积加到一起为 40+60=100 英里, 也就是路程.

让我们再重新考虑第二个长方形. 假设我们说它的 "高" 为 -30(英里/小时), 因为该长方形在横坐标轴的下方. 当然, 一个长方形的高实际上是不可能为负的, 但无论如何区分坐标轴上和下是很必要的. 所以如果它的高度为 -30, 面积为 $2\times(-30) = -60$ 英里. 让我们把这个负号去掉, 正确地把它标记为有向面积. 我们的结论是：在坐标轴下方的面积为负. 如果这样做, 总面积是 40(第一个长方形) 加上 -60(第二个长方形), 得到的面积为 -20. 我们刚才计算的位移不就是 -20 嘛!

在上一节的公式中, 我们对时间段 $[3, 6]$ 有一个分区 $3 < 4 < 6$. 第一个分区的速度为 $v_1 = 40$, 第二分区的速度为 $v_2 = -30$. 所以有

$$\begin{aligned}\text{位移} = \sum_{j=1}^{n} v_j(t_j - t_{j-1}) &= v_1(t_1 - t_0) + v_2(t_2 - t_1) \\ &= 40 \times (4-3) + (-30) \times (6-4) = -20.\end{aligned}$$

在第二时间段内, 如果说速率 (而不是速度) $v_2 = 30$, 那么总和为 $40 \times (4-3) + 30 \times (6-4) = 100$, 这就是我们刚才计算出的路程. 当然, 速率为 30 英里/小时是速度为 -30 英里/小时的绝对值. 所以如果不用没有方向的面积去计算路程, 我们可以用速度的绝对值 $|v|$ 对时间 t 的图像来表示, 如图 15-6 所示.

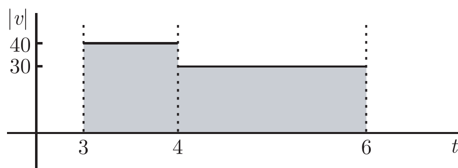


图 15-6

现在, 面积的方向已经不是很重要了, 因为坐标轴以下没有面积了! 所以我们可以说所有的面积都是有向的. 如果我们考虑有向面积, 应该先取绝对值. 详情参见 16.4.1 节.

15.2.4 连续的速度

我们已经看到如果一辆车 (或一个物体) 沿直线行驶, 它的速度在时间 $[a, b]$ 内的有限时间段 (分区) 内是一个常数, 这时位移为速度对时间的图像与 t 轴及 $t = a$ 和 $t = b$ 所围成的有向面积. 对于路程也是同样的, 唯一的区别是, 这时的图像是速度的绝对值 $|v|$ 对时间 t 的图像.

如果在有限时间分段内的速度不是常数, 那情况又是如何呢? 除非你从来不关闭控制系统, 否则可能会不时地为超车而加速或当看到警察时减速 …… 你的速度从 40 英里/小时到 60 英里/小时需要一定的加速, 因为你不可能一下子加速. 所

以, 让我们考虑速度为时间 t 的**连续函数**, 如图 15-7 所示.

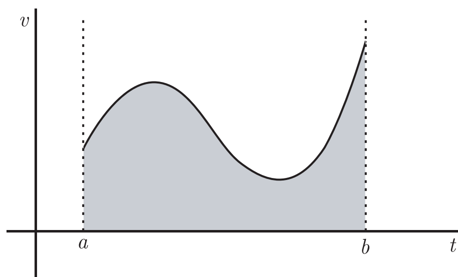


图 15-7

该车加速, 然后减速, 最后又更快地加速. 位移应该是阴影部分的面积 —— 实际上阴影部分都是在坐标轴上方, 所以该位移也是路程. 我们究竟怎样计算这块面积呢?

下面是我们的想法. 在从 a 到 b 的时间段内, 速度变化了很多, 但在一个非常小的时间段内, 速度并没有发生很大的变化. 让我们考虑一个小的时间段, 叫它为 $[p, q]$, 我们研究在这个小时间段内的情况. 即使在这个小时间段内, 速度也是有微小变化的, 但我们假设速度没变. 我们在时间段 $[p, q]$ 内选择某一时刻 c 的速度作为样本速度, 看看这时的速度为多少. 我们也假设所选择的样本速度为该时间段内 $[p, q]$ 的实际速度. 如果把速度 v 写为 $v(t)$ 来强调速度 v 为时间 t 的函数, 这时在时刻 c 的速度就为 $v(c)$. 所以, 我们有图 15-8.

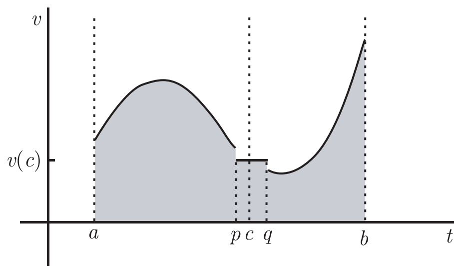


图 15-8

对于 $[p, q]$ 这个时间段内的图像, 我们已经把它以高度 $v(c)$ 画平. 这样做的好处是, 可以求出在 $[p, q]$ 这个时间段内的位移. 这块小长方形的高为 $v(c)$, 底为 $q-p$, 所以它的面积为 $v(c)\times(q-p)$. 这实际上不是那段时间内的准确位移, 但相当接近.

讨论为什么要止步于这个小区间 $[p, q]$ 呢? 让我们在整个 $[a, b]$ 区间内重复这个划分的过程. 以下面这个分区开始

$$a=t_0<t_1<t_2<\cdots<t_{n-2}<t_{n-1}<t_n=b,$$

在每个时间段内, 我们都取个样本速度. 第一个时间段是从 t_0 到 t_1, 所以我们选在那个时间段的某一时刻 c_1, 假设在这个时间段内的速度为 $v(c_1)$. c_1 这个数可能等于该时刻的开始值 t_0 或结束值 t_1, 或在该时间段内的任意值, 无论什么值, 只要它在 $[t_0, t_1]$ 这个时间段内. 现在, 对第二个时间段重复这个过程. 在 $[t_1, t_2]$ 内, 我们选择 c_2, $v(c_2)$ 作为这个时间段内的样本速度. 对以后的每个时间段用同样的方法, 直到在 $[t_{n-1}, t_n]$ 这个时间段内我们选 c_n. 图 15-9 是当 n =6 时的例子.

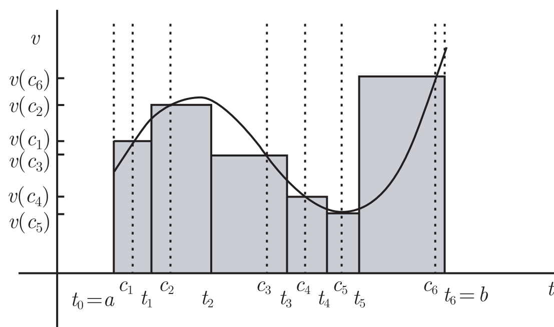


图 15-9

到目前为止, 我们所做的是使用像这些楼梯一样的函数 (其中的每一级台阶都与这个函数有交点) 去逼近这条平滑的速度曲线. 我们可以使用上一节的一些方法计算阴影部分的面积, 但这个计算仅仅是对实际面积的一个估算. 我们得到

$$\text{速度曲线下的面积} \approx \sum_{j=1}^{n} v(c_j)(t_j - t_{j-1}).$$

不幸的是, 我们的估算不是很准确. 图像右侧的大长方形对于 $[t_5, t_6]$ 区间内的曲线下面积的估算不是很准确, 因为在曲线上部有太多的长方形面积. 所以让我们重新划分更多的区, 把每个分区的区间缩小, 如图 15-10 所示.

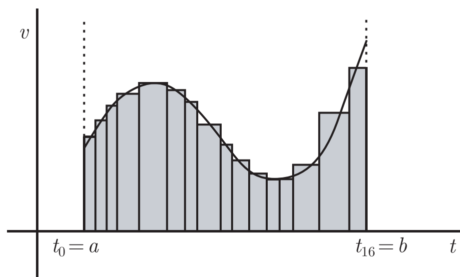


图 15-10

在这个图像中, 我们有 16 个分区而不是 6 个了, 看起来, 阴影部分的面积比之前的分区更接近于真实面积了. 尽管我们在分区中可以使用很多小区间, 但如果其

中的某个分区很大, 对估算结果仍然会有很大的影响, 如图 15-11 所示.

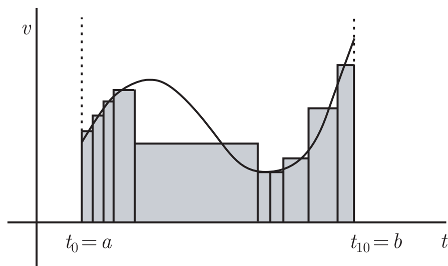


图 15-11

即使大多数的长方形的宽都很小, 但只要其中有一个长方形的宽很大, 就会严重影响计算结果. 所以, 需要其中的每一个分区间隔都很小. 我们把其中最大的间隔叫**最大区间**, 解决这个问题的方法是让**最大区间**足够小, 最终它的极限为 0. 用这种方式, 我们可以说所有的分区间隔都会很小, 再也不会出现刚才图像的那种情况了.

正式的说法是, 这个最大区间可以被定义为

最大区间 $=(t_1 - t_0), (t_2 - t_1), \cdots, (t_{n-1} - t_{n-2}), (t_n - t_{n-1})$ 的最大值.

例如, 如果在 [3, 5] 区间内的划分是 $3 < 3.25 < 4 < 4.5 < 5$ (这是在 15.2.1 节中对第三辆车已经使用过的划分), 这时, 这些小分区的长度分别为 0.25(3.25 − 3), 0.75 (来自于 4 − 3.25), 0.5(4.5 − 4) 和 0.5(5 − 4.5). 在 0.25, 0.75, 0.5 和 0.5 中的最大值是 0.75, 所以这个最大区间是 0.75.



现在, 我们用极限的方法来替代估算

$$\sum_{j=1}^{n} v(c_j)(t_j - t_{j-1}),$$

从而得到实际的数值. 假设我们不断重复上述过程, 每一次都确保这次的最大区间比上一次的要小, 所以这个最大值最终趋于 0. 这样, 这个估算越来越精确了. 这就是我们尽量想得到的公式:

$$\text{在速度曲线下的实际面积} = \lim_{\text{mesh}\to 0} \sum_{j=1}^{n} v(c_j)(t_j - t_{j-1}).$$

因为最大区间趋于 0, 这样划分的数目就会越来越大, 所以上述极限自动包含了 $n \to \infty$ 这样一个思想.

15.2.5 两个特别的估算

上述的公式还有许多待改进之处. 如果我们选择不同的划分, 使用不同的样本时间 c_j, 还会得到同样的答案吗? 这实际上是一个定理, 如果 v 是关于时间 t 的连

续函数, 这时上述极限是独立于划分和样本时间的. 定理的证明不属于本书的范围, 但是在大多数的数学分析教材中都能找到. 另一方面, 我们可以通过两个特例来感受这个定理的基本思想：上和与下和.

从一个划分开始, 我们在每一个小分区中选一些样本点. 假设我们总是选一个点对应其所在区间的可能的最大速度. 例如, 在区间 $[t_0, t_1]$ 中选点 c_1 使得 $v(c_1)$ 是速度 v 在这个时间段的最大值. 对于每个时间段, 我们都做同样的取值. 这说明取值在曲线之上. 图 15-12 是这种情况的图像.

图 15-12 中的长方形面积, 我们叫作**上和**, 很明显, 这比实际面积要大. 另一方面, 如果我们对于每一个分区都选择最小的速度, 将得到图 15-13.

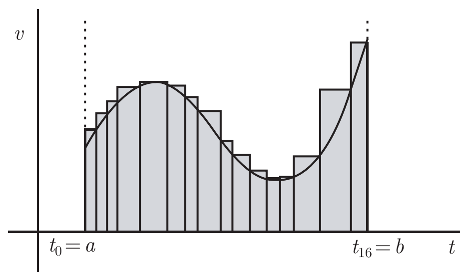


图 15-12

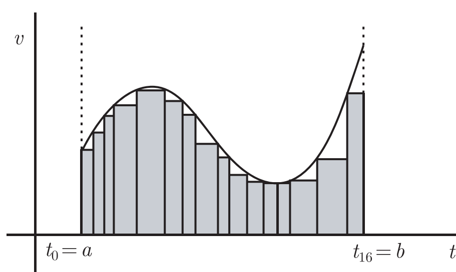


图 15-13

划分是同样的, 但样本时间不同. 由于我们所选取的方式不同, 这次所有的长方形都在曲线的下方; 这样的所有长方形的面积和叫作**下和**, 它比实际的面积要小.

通过对这两种情况的分析, 我们有

$$\text{下和} \leqslant \text{曲线下的实际面积} \leqslant \text{上和}$$

实际上, 对于同样的划分, 无论我们选什么样的样本时间 c_j, 它所对应的长方形面积都在上和与下和的面积之间. 如果对于每一个分区我们都考虑使它的最大区间足够小, 这时上和与下和的极限值就会是一样的 (但我并不打算证明这个理论). 之前学过的三明治定理将会证明这个公式是有意义的. 无论选取怎样的 c_j 值, 你的求和都是在上和与下和之间. 当最大区间趋于 0 时, 三明治定理可以证明这两个求和都趋于实际的正确面积.

现在, 我们有了所有需要定义定积分的工具. 现在就开始讨论它 ……

第 16 章　定　积　分

现在该介绍定积分划分了. 首先我们用面积来给定积分下一个非正式的定义, 接下来使用上一章的划分思想来使这个定义更严谨. 在学习了一个应用严格定义的 (实际很烦琐的) 例子后, 我们将会对定积分的定义有进一步的理解. 更准确地说, 我们将会学到以下知识点:

- 有向面积和定积分;
- 定积分的定义;
- 使用这个定义的例子;
- 定积分的基本性质;
- 使用积分求解面积 —— 两条曲线之间的面积, 以及在一条曲线和 y 轴之间的面积;
- 估算定积分;
- 函数的平均值和定积分的中值定理;
- 一个不可积函数的例子.

16.1　基 本 思 想

我们从一个函数以及 $[a,b]$ 区间开始研究. 画出 $y=f(x)$ 这个函数图像, 考虑该曲线, x 轴和两条垂直线 $x=a$ 和 $x=b$ 所围成的面积 (如图 16-1 所示).

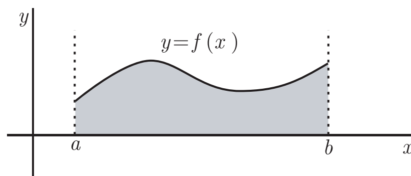


图　16-1

能有一种简洁的方式来表示阴影部分的面积就太好了. 因为上述图像中没有长度单位, 所以我们将用 “单位” 度量长度, 用 “平方单位” 度量面积. (如果上述图像有单位, 比如英寸, 那么它的面积单位会是平方英寸.) 无论如何, 我们都可以说阴影部分的面积 (平方单位) 是

$$\int_a^b f(x)\mathrm{d}x.$$

这就是**定积分**. 你可以把它读为 “函数 $f(x)$ 对于 x 从 a 到 b 的积分”. 表达式 $f(x)$

叫作**被积函数**, 它告诉你这条曲线是什么样子. a 和 b 说明两条垂线在哪, 也叫**积分极限**(请注意不要同极限混在一起!) 或**积分端点**. 最后, $\mathrm{d}x$ 说明 x 是水平轴的变量. 实际上, x 是虚拟变量 —— 你可以用任意其他字母来表示它. 所以下列表达式是等价的：

$$\int_a^b f(x)\mathrm{d}x = \int_a^b f(t)\mathrm{d}t = \int_a^b f(q)\mathrm{d}q = \int_a^b f(\beta)\mathrm{d}\beta.$$

实际上, 这些表达式的计算结果是相同的, 都是上述图像阴影部分的面积 (平方单位); 不同的仅仅是, 我们把横坐标从 x 轴更名为 t 轴、q 轴或 β 轴. 但这并不影响对面积的计算!

如果函数有一部分在 x 轴的下方, 情况又会怎样？图像可能看起来如图 16-2 所示.

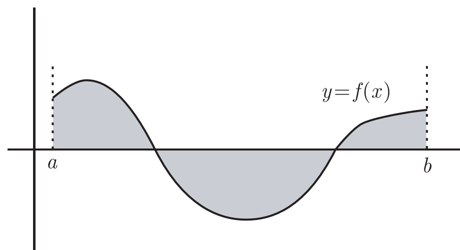


图　16-2

像我们在 15.2.3 节中看到的, 只有把 x 轴下方的面积作为负面积来看时, 才有意义. 如果在 $x=a$ 和 $x=b$ 之间的曲线的所有部分都在坐标轴下方, 那么该积分一定为负的. 实际上, 该积分给出了有向面积的大小. 更准确地表述如下.

> $\int_a^b f(x)\mathrm{d}x$ 是由曲线 $y=f(x)$, 两条垂线 $x=a$ 和 $x=b$, 以及 x 轴所围成的有向面积(平方单位).

注意积分只是个数, 但面积是有单位的.

从上一章中我们知道, 在时间 a 和 b 之间的一个物体的位移是两条垂线 $t=a$ 和 $t=b$, t 轴以及曲线 $y=v(t)$ 所围成的有向面积. 路程同位移的计算方法基本相似, 但有一点不同 (很关键), 那就是 $y=|v(t)|$. 用我们的符号可以表示如下.

$$\boxed{位移 = \int_a^b v(t)\mathrm{d}t} \quad 和 \quad \boxed{路程 = \int_a^b |v(t)|\mathrm{d}t.}$$

对这个问题的理解是：从 $t=a$ 开始, 以 $t=b$ 结束. 注意该问题的虚拟变量是 t, 被积函数分别是速度 $v(t)$ 和速率 $|v(t)|$.

一些简单的例子

现在来看一些关于定积分的简单例子. 首先, 考虑

$$\int_0^1 x\mathrm{d}x \quad 和 \quad \int_0^2 x\mathrm{d}x.$$

两道例题的被积函数都是 x, 所以我们从绘制 $y = x$ 的函数图像开始. 前个例子的面积从 x =0 到 x =1, 而后个例子的面积从 x =0 到 x =2. 我们看到如图 16-3 的两个面积.

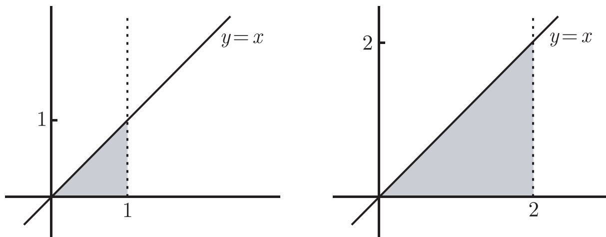


图 16-3

这些面积很容易计算：两个都是三角形. 第一个三角形的底和高都是 1, 所以面积是 $\frac{1}{2} \times (1) \times (1) = \frac{1}{2}$ 平方单位; 第二个三角形的底和高都是 2, 所以面积为 $\frac{1}{2} \times (2) \times (2) = 2$ 平方单位. 表示为

$$\int_0^1 x\mathrm{d}x = \frac{1}{2} \quad 和 \quad \int_0^2 x\mathrm{d}x = 2.$$

现在我们使用这些公式解决实际问题. 假设一辆车开始启动, 加速度为常数 1 码/二次方秒; 它的速度 (码/秒) 为 $v(t) = t$. 一秒钟之后该车走了多远？两秒钟之后呢？答案已由上边的积分给出了. 仅仅用 t 替代 x, 你就会得到答案. 首先, 注意对于这个问题位移和距离是同样的, 因为这辆车一直沿正方向行驶. 所以在第一秒, 我们有

$$位移 = \int_0^1 v(t)\mathrm{d}t = \int_0^1 t\mathrm{d}t = \frac{1}{2}.$$

前两秒有

$$位移 = \int_0^2 v(t)\mathrm{d}t = \int_0^2 t\mathrm{d}t = 2.$$

当然, 这些位移是以码为单位的.

现在来看另一个定积分

$$\int_{-2}^5 1\mathrm{d}x.$$

为求这个定积分的值, 我们需要绘制函数 y =1 的图像, 位于垂线 $x = -2$ 和 x =5 之间. 想要计算的面积如图 16-4 所示.



图 16-4

它所围成的面积为长方形, 高为 1, 底为 7, 面积为 7 平方单位. 这就是说

$$\int_{-2}^{5} 1\mathrm{d}x = 7.$$

事实上, 这个一般的积分表达式

$$\int_a^b 1\mathrm{d}x$$

表示的面积如图 16-5 所示.

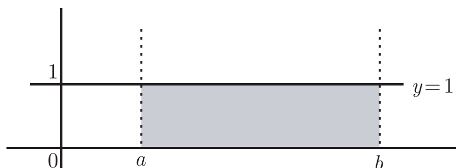


图 16-5

该长方形的高为 1, 底边长为 $b-a$ (即使 a 和 b 是负的), 所以我们有了一般的表达方式:

$$\int_a^b 1\mathrm{d}x = b-a,$$

也可以简单地写为

$$\int_a^b \mathrm{d}x = b-a.$$

因为我们可以认为 $1\mathrm{d}x$ 就是 $\mathrm{d}x$.

式子

$$\int_{-\pi}^{\pi} \sin(x)\mathrm{d}x$$

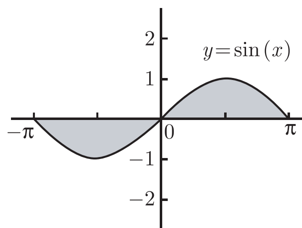


图 16-6

表示什么呢? 让我们绘制函数图像, 看看将要计算的面积是什么样的, 如图 16-6 所示.

幸运的是, 我们要计算的是有向面积, 而不是实际的面积. 根据对称性, x 轴上方 (在 0 和 π 之间) 的面积同 x 轴下方 (在 −π 和 0 之间) 的面积是一样的. 所以正负面积互相抵消, 最后求得的总面积为 0 平方单位. 也就是,

$$\int_{-\pi}^{\pi} \sin(x)\mathrm{d}x = 0.$$

如果你想求得**实际**的面积, 而不是有向面积, 需要仔细考虑把积分分为两部分来计算. 在 16.4.1 节中, 我们将讲到如何计算, 在 17.6.3 节中, 你会看到同样的例题.

在看下一道例题之前, 我希望对刚才的例题做个总结. 上述积分为 0 的理由是：被积函数 $\sin(x)$ 为奇函数, 被积区间 $[-\pi, \pi]$ 是关于原点对称的. 我们可以用其他的任意奇函数替代 $\sin(x)$, 把积分区间改为从 $-a$ 到 a(a 为任意数), 积分结果仍然为 0. 也就是说,



$$\text{如果 } f \text{ 为奇函数}, \int_{-a}^{a} f(x)\mathrm{d}x = 0 \quad (a \text{ 为任意实数}).$$

根据对称性可知：在 x 轴上方的任意一块面积都可以找到其在 x 轴下方的对应面积, 就像刚才图中表示的那样. 如果要计算的积分符合上述条件, 那么就没有必要做计算了, 这会节省我们很多时间. 在 18.1.1 节中, 我们将会给出这个结论的正式证明.

16.2 定积分的定义

关于用面积给出的定积分的定义, 我们已经给出了很多的例子, 但这并不能帮助我们解决如何计算一些特殊的积分. 确实, 上一节的每一个例子, 我们都找到了答案, 这仅仅是因为我们知道怎样计算三角形或长方形的面积. 更幸运的是最后一个例子 $\sin(x)$, 因为两个面积被抵消了. 但在通常情况下, 我们没有这么幸运.

事实上, 在以前的导数学习中, 我们遇到过这种情况. 我们已经定义了导数 $f'(x)$ 的几何意义是函数 $y = f(x)$ 在点 $(x, f(x))$ 的切线的斜率, 但并不知道如何求得斜率. 取而代之, 我们定义导数为

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

假设该极限是存在的. 像我们从前观察过的一样, 该极限是 0/0 型不定式, 但在很多情形下我们可以计算出该极限. 无论如何, 一旦给出上述定义, 那么 $f'(x)$ 就可以表示切线的斜率.

遗憾的是, 定积分的定义没这么简单, 它比导数的定义要复杂得多. 好在我们已经在前一章做了很多工作, 可以把它定义如下.

$$\int_{a}^{b} f(x)\mathrm{d}x = \lim_{\mathrm{mesh} \to 0} \sum_{j=1}^{n} f(c_j)\,(x_j - x_{j-1}),$$

其中 $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ 并且对于每一个 $j = 1, \cdots, n$ 都有 c_j 在 $[x_{j-1}, x_j]$ 内.

这个定义尽管很长, 仍然没有告诉我们全部内容! 你仍然需要注意如下几点.

- 表达式 $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ 告诉我们, 点 $x_0, x_1, x_2, \cdots, x_{n-1}, x_n$ 形成了区间 $[a, b]$ 的划分, 其中最左边的 $x_0 = a$, 最右边的 $x_n = b$. 这个划分创造了 n 个小的子区间 $[x_0, x_1], [x_1, x_2], \cdots, [x_{n-1}, x_n]$.
- 划分中的最大区间是指所有这些小区间中最长的区间, 所以我们有:
  $\text{mesh} = (x_1 - x_0), (x_2 - x_1), \cdots, (x_{n-1} - x_{n-2}), (x_n - x_{n-1})$ 中的最长区间.
- 对于每一个小区间, c_j 可以被选择在它所对应区间的任何位置. 这就是我们为什么说 c_j 在 $[x_{j-1}, x_j]$ 区间内.
- 上述极限是不断计算最大区间越来越少的不同的划分而求得的; 也就是说, 当它的最大区间趋于 0 时, 我们会有 n 趋于无穷大. 每一个划分都涉及 c_j 的选择.
- 如果 f 是连续的函数, 那我们怎样划分以及怎样选择 c_j 就显得无关紧要了, 只要它的最大区间趋于 0. 事实上, 只要函数 f 是有界的, 即使它有有限个不连续的点, 这也是成立的. 这样的函数是**可积的**, 因为它可被积分. 也有一些函数, 即使它有无穷多个不连续的点, 也是可积的, 但这已经超出了本书的讨论范围. 另一方面, 如果函数 f 是无界的, 也可能是可积的, 比如它有垂直渐近线, 这种积分叫作反常积分, 参见第 20 章和第 21 章对这个问题的讲解.
- 在积分表达式中出现的求和 $\sum_{j=1}^{n} f(c_j)(x_j - x_{j-1})$, 我们称之为**黎曼和**. 它给出了定积分的估算值. 如果它的最大区间非常小, 那么这时估算将是非常精确的.

看到了吧, 我说过这很复杂的! 现在, 我们看看怎样用这个定义计算定积分.

一个使用定义的例子

我们来看如何用上述公式计算定积分

$$\int_0^2 x^2 \mathrm{d}x.$$

我们看图 16-7.

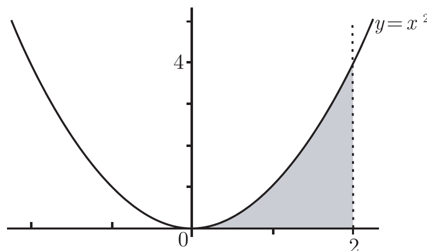


图　16-7

这不是三角形, 也不是长方形, 它的面积都在 x 轴上方, 所以也没有可抵消的部分. 我们设 $f(x)=x^2$, 使用定积分的定义计算面积.

我们需要用最大区间越来越小的划分来解决这个问题. 到目前为止, 最简单的方式是用大小相等的小区间解决问题. 所以需要把 [0,2] 区间分成 n 个小区间, 每个小区间的长度是相等的. 因为总长度为 2, 共有 n 个区间, 所以每个区间的长度为 $2/n$. 第一区间是从 0 到 $2/n$; 第二区间是从 $2/n$ 到 $4/n$, 以此类推. 把图 16-7 中的阴影部分放大, 我们得到图 16-8.

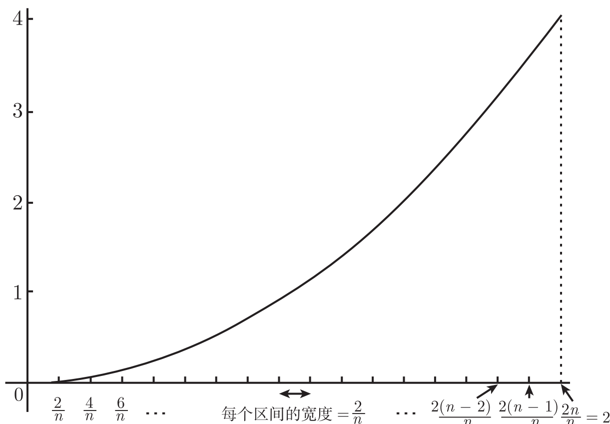


图 16-8

在该例子中, 通常的分区

$$a=x_0<x_1<x_2<\cdots<x_{n-1}<x_n=b$$

特殊化为

$$0=\frac{0}{n}<\frac{2}{n}<\frac{4}{n}<\cdots<\frac{2(n-1)}{n}<\frac{2n}{n}=2.$$

该划分的最大区间的长度为 $2/n$, 其实每个区间的长度都为 $2/n$. 很显然, 对于任意一点 x_j, 它所对应的横坐标为 $2j/n$. 现在我们需要选择 c_j. 例如 c_0 可能在区间 $[0,2/n]$ 的任意位置, c_1 可能在 $[2/n, 4/n]$ 的任意位置, 以此类推. 为了计算简单, 我们可以选择每一个小区间的右端点, 所以 $c_j=x_j=2j/n$. 因此, $c_j=\dfrac{2j}{n}$ 是在小区间 $[x_{j-1},x_j]=\left[\dfrac{2(j-1)}{n},\dfrac{2j}{n}\right]$ 上的选择.

这样, 我们有如图 16-9 所示的一系列长方形.

我们正在计算的是上和 —— 因为所有的长方形都有一部分在曲线的上面. (参见 15.2.5 节中关于上和的讨论.)

现在, 我们准备使用公式了. 考虑黎曼和

$$\sum_{j=1}^{n} f(c_j)(x_j - x_{j-1}).$$

我们知道 $f(x)=x^2, c_j=2j/n, x_j=2j/n$, $x_{j-1}=2(j-1)/n$, 所以上述求和变为

$$\sum_{j=1}^{n}\left(\frac{2j}{n}\right)^2\left(\frac{2j}{n}-\frac{2(j-1)}{n}\right).$$

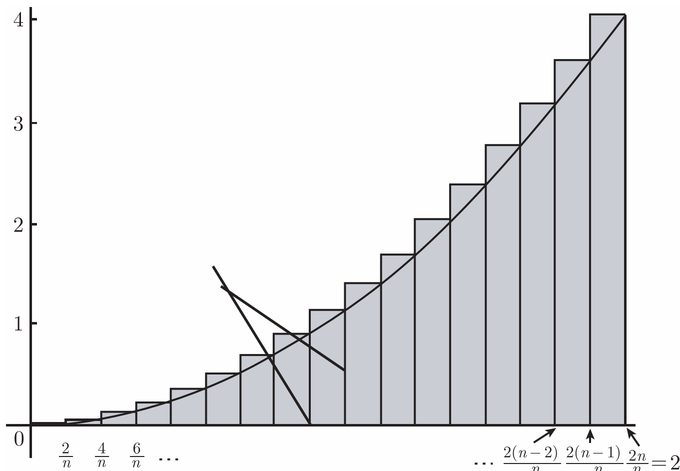


图 16-9

等式右侧括号中的差化简后为 $2/n$. 我们对这个结果并不感到惊奇, 因为这就是每个长方形的宽. 另外, 这些长方形的宽虽相同, 但高却不同, 第 j 个长方形的高为 $(2j/n)^2$, j 的取值范围从 1 到 n. 这样, 上述的和可化简为

$$\sum_{j=1}^{n}\frac{4j^2}{n^2}\times\frac{2}{n}=\sum_{j=1}^{n}\frac{8j^2}{n^3}.$$

很好, 我们发现分母 n^3 与虚拟变量 j 无关, 所以可以把它当作公因子提出来, 这样该求和表达式为

$$\frac{8}{n^3}\sum_{j=1}^{n}j^2.$$

在 15.1.2 节中, 我们已经知道上述和的值为 $n(n+1)(2n+1)/6$. 也就是说,

$$\frac{8}{n^3}\sum_{j=1}^{n}j^2=\frac{8}{n^3}\times\frac{n(n+1)(2n+1)}{6}=\frac{4(n+1)(2n+1)}{3n^2}.$$

最终, 通过计算可知, 刚才图像中阴影部分的面积为

$$\sum_{j=1}^{n} f(c_j)(x_j - x_{j-1})=\frac{4(n+1)(2n+1)}{3n^2}.$$

这仅仅是对阴影部分面积的一个估算. 因为该划分的每个区间的宽度都为 $2/n$, 所以可以通过让 $n \to \infty$ 而迫使 $2/n$ 趋于 0. 这样长方形变得越来越小, 个数越来越多, 我们的计算也就越来越精确了. 所以我们有

$$\int_0^2 x^2 \mathrm{d}x = \lim_{\text{最大区间} \to 0} \sum_{j=1}^{n} f(c_j)(x_j - x_{j-1}) = \lim_{n \to \infty} \frac{4(n+1)(2n+1)}{3n^2}.$$

这样余下要做的是求解这个极限. 我们可以使用 4.3 节的方法来求解这个极限, 该极限的结果为 8/3, 所以最后的结论为

$$\int_0^2 x^2 \mathrm{d}x = \frac{8}{3}.$$



这样算出的面积为 8/3 平方单位. 现在, 请你使用刚才介绍的方法证明

$$\int_0^1 x^2 \mathrm{d}x = \frac{1}{3}.$$

像刚才计算的那样, 这个方法很烦琐. 不仅因为这个计算很长, 也因为我们需要知道怎样求下面这个和

$$\sum_{j=1}^{n} j^2.$$

如果被积函数是 x^3 而不是 x^2, 那么我们就需要计算和式

$$\sum_{j=1}^{n} j^3.$$

但如果被积函数为 $\sin(x)$ 或其他类似函数, 情况会变得很糟糕. 所以我们需要一个不用长方形和求和的方法. 但这要等到下一章讲了微积分的第二基本定理才能找到答案. 接下来, 我们看看定积分都有什么性质.

16.3 定积分的性质



我们再将定积分的定义扩展些. 你对

$$\int_2^0 x^2 \mathrm{d}x$$

这个定积分怎样看?

这个积分同我们上一节计算过的积分的唯一不同是, 它是从 2 到 0 而不是从 0 到 2. 所以怎样划分 [2, 0] 呢? 这并不是一个正常的区间, 因为 2 比 0 大. 最好的解决方式是采用同刚才的划分相反的方式, 如下所示:

$$2 = x_0 > x_1 > x_2 > \cdots > x_{n-1} > x_n = 0.$$

现在, 在上述定义中出现的值 $(x_j - x_{j-1})$ 总是为负的. 实际上, 这个长方形的底长为负的! 这样该积分的结果为

$$\int_2^0 x^2 \mathrm{d}x = -\frac{8}{3}.$$

所以, 如果翻转积分, 即调换积分上下限, 需要在这个积分前面加个负号. 总的来说, 对于可积函数 f 以及常数 a 和 b, 我们有

$$\boxed{\int_b^a f(x)\mathrm{d}x = -\int_a^b f(x)\mathrm{d}x.}$$

这个公式的另一个解释是：对于一个正在做直线运动的物体, 考虑该物体向回走的情况, 这时的位移就是负的了. 例如, 你拍摄汽车前进的影片, 然后把影片回放, 那么汽车就是倒着开的, 这时的位移是负的.

现在如果积分上下限是相等的, 那结果又是怎样的? 例如, 考虑

$$\int_3^3 x^2 \mathrm{d}x.$$

这并不是一个面积. 毕竟, 在 x =3 和 x =3 之间没有面积. 所以答案是 0. 实际上, 对于任意实数 a, 都有

$$\boxed{\int_a^a f(x)\mathrm{d}x = 0}$$

我们可以再一次用物理学的直线运动来解释：在时间 a 和 a 之间, 实际上根本就没有时间, 物体也不可能移动, 所以根本就没有位移.

接下来, 我们考虑这个图像, 如图 16-10 所示.

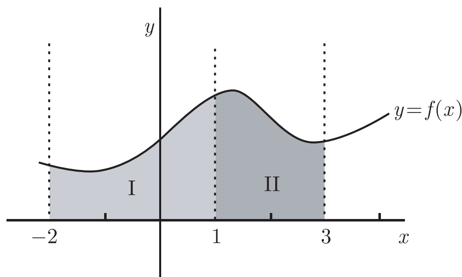


图　16-10

从 $x = -2$ 到 $x = 3$ 的整个面积, 很显然是 I 和 II 两部分的面积和. 通过定义, 我们分别有

$$\text{面积 I} = \int_{-2}^1 f(x)\mathrm{d}x \quad \text{和} \quad \text{面积 II} = \int_1^3 f(x)\mathrm{d}x,$$

这样可以得到结论

$$\int_{-2}^{3} f(x)\mathrm{d}x = \int_{-2}^{1} f(x)\mathrm{d}x + \int_{1}^{3} f(x)\mathrm{d}x.$$

我们所要做的就是把这个面积分成两部分, 然后分别用积分来表示. 当然, 我们也可以用在区间 [−2,3] 上的任意数来拆开这个积分, 只要我们用同一个数来替代两个积分表达式中的 1. 事实上, 即使我们选的数不在 [−2, 3] 区间, 这个划分方法也是适用的. 例如, 下述这个公式是正确的:

$$\int_{-2}^{3} f(x)\mathrm{d}x = \int_{-2}^{4} f(x)\mathrm{d}x + \int_{4}^{3} f(x)\mathrm{d}x.$$

图 16-11 是对应这个公式的图像.

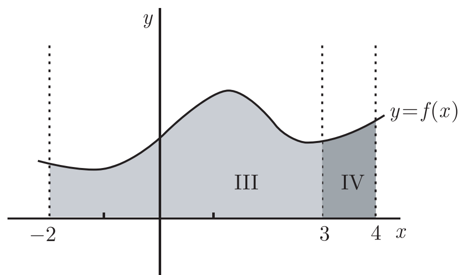


图 16-11

这时, 我们有

$$面积\mathrm{III} = \int_{-2}^{3} f(x)\mathrm{d}x \quad 和 \quad 面积\mathrm{IV} = \int_{3}^{4} f(x)\mathrm{d}x.$$

把这两个积分加到一起就有

$$\int_{-2}^{4} f(x)\mathrm{d}x = \int_{-2}^{3} f(x)\mathrm{d}x + \int_{3}^{4} f(x)\mathrm{d}x.$$

现在把这个等式最右侧的积分表达式的积分上下限互换:

$$\int_{-2}^{4} f(x)\mathrm{d}x = \int_{-2}^{3} f(x)\mathrm{d}x - \int_{4}^{3} f(x)\mathrm{d}x.$$

整理这个等式, 就得到了我们想要的等式

$$\int_{-2}^{3} f(x)\mathrm{d}x = \int_{-2}^{4} f(x)\mathrm{d}x + \int_{4}^{3} f(x)\mathrm{d}x.$$

总的来说, 对于任何可积函数 f 以及常数 a、b、c, 我们都有

$$\boxed{\int_{a}^{b} f(x)\mathrm{d}x = \int_{a}^{c} f(x)\mathrm{d}x + \int_{c}^{b} f(x)\mathrm{d}x.}$$

我们可以把一个积分表达式分成两部分, 即使分隔点 c 是在原始区间 $[a, b]$ 之外, 当然要求分隔之后的两部分依然是可积的.



例如, 求

$$\int_1^2 x^2 \mathrm{d}x,$$

我们可以使用上节已经得到的结论

$$\int_0^2 x^2 \mathrm{d}x = \frac{8}{3} \quad \text{和} \quad \int_0^1 x^2 \mathrm{d}x = \frac{1}{3}.$$

你需要做的是把第一个积分在 $x = 1$ 处分成两部分, 像这样:

$$\int_0^2 x^2 \mathrm{d}x = \int_0^1 x^2 \mathrm{d}x + \int_1^2 x^2 \mathrm{d}x.$$

使用上述结论, 我们有

$$\frac{8}{3} = \frac{1}{3} + \int_1^2 x^2 \mathrm{d}x,$$

这样, 结果为

$$\int_1^2 x^2 \mathrm{d}x = \frac{8}{3} - \frac{1}{3} = \frac{7}{3}.$$



现在留给你去证明这个:

$$\int_1^2 x \mathrm{d}x = \frac{3}{2}.$$

你可以使用 16.1 节得到的结论:

$$\int_0^2 x \mathrm{d}x = 2 \quad \text{和} \quad \int_0^1 x \mathrm{d}x = \frac{1}{2}.$$

这里还有两个更简单却很实用的积分性质. 首先是**常数可以被移到积分表达式的外边**. 也就是说, 对于任何可积函数 f, 常数 a、b、C, 都有

$$\boxed{\int_a^b Cf(x) \mathrm{d}x = C \int_a^b f(x) \mathrm{d}x.}$$

如果 C 是个关于 x 的函数, 那么该表达式就不成立了! C 一定要是一个常数. 实际上, 证明这个很容易. 只要写为

$$\int_a^b Cf(x) \mathrm{d}x = \lim_{\text{最大区间} \to 0} \sum_{j=1}^{n} Cf(c_j)(x_j - x_{j-1})$$

把常数 C 移到求和符号的外边, 这时极限为

$$\int_a^b Cf(x) \mathrm{d}x = C \lim_{\text{最大区间} \to 0} \sum_{j=1}^{n} f(c_j)(x_j - x_{j-1}) = C \int_a^b f(x) \mathrm{d}x.$$

例如, 求

$$\int_0^2 7x^2\mathrm{d}x.$$

只要把 7 移出积分符号即可：

$$\int_0^2 7x^2\mathrm{d}x = 7\int_0^2 x^2\mathrm{d}x = 7\times\left(\frac{8}{3}\right) = \frac{56}{3}.$$

第二个性质是**和或差的积分等于积分的和或差**. 也就是说, 如果 f 和 g 都为可积函数, a 和 b 为常数, 这时

$$\boxed{\int_a^b (f(x)+g(x))\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x.}$$

如果把加号变为减号也是成立的. 无论是加还是减, 用拆分法证明都是很容易的. 例如对于加法要做的是把和写成极限的形式, 像这样：

$$\begin{aligned}\int_a^b (f(x)+g(x))\mathrm{d}x &= \lim_{\text{最大区间}\to 0}\sum_{j=1}^{n}(f(c_j)+g(c_j))(x_j-x_{j-1})\\ &= \lim_{\text{最大区间}\to 0}\sum_{j=1}^{n} f(c_j)(x_j-x_{j-1}) + \lim_{\text{最大区间}\to 0}\sum_{j=1}^{n} g(c_j)(x_j-x_{j-1})\\ &= \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x.\end{aligned}$$

可以用同样的方法证明减法时依然成立.

例如, 求

$$\int_0^2 (3x^2-5x)\mathrm{d}x,$$

把这个积分分成两部分, 同时把常数提出来, 这时有

$$\int_0^2 (3x^2-5x)\mathrm{d}x = 3\int_0^2 x^2\mathrm{d}x - 5\int_0^2 x\mathrm{d}x = 3\times\left(\frac{8}{3}\right) - 5\times(2) = -2.$$

这里我们使用了刚才的结论

$$\int_0^2 x^2\mathrm{d}x = \frac{8}{3} \quad \text{和} \quad \int_0^2 x\mathrm{d}x = 2.$$

16.4 求 面 积

如果 $y=f(x)$, 我们可以不用 $f(x)$ 作为被积函数, 而把它写成 $\int_a^b y\mathrm{d}x$. 这个表达式有个很好的几何解释：通过划分方法, 我们观察其中的一个小长方形, 也可以说是一个小竖条, 它的高为 y 单位, 宽很小为 $\mathrm{d}x$ 单位 (如图 16-12 所示).

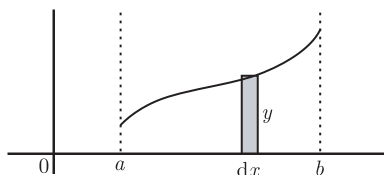


图　16-12

该小竖条的面积为高乘以宽, 即 $y\mathrm{d}x$ 平方单位. 现在我们划分更多的竖条, 划分 $[a,b]$ 区间. 如果把所有竖条的面积加到一起, 我们将得到该面积的一个近似值. 这个积分符号不仅是求和的意思, 同时它也要求所有竖条 (即长方形) 的宽度趋于 0(极限的方法).

这个思想很关键, 会帮助我们理解怎样用积分计算面积. 现在让我们花一些时间去看看三种特殊的面积：通常的面积, 两条曲线之间的面积, 曲线和 y 轴所围成的面积.

16.4.1　求通常的面积



我们已经知道, 定积分所处理的是有向面积. 很显然, 如果曲线一直是在 x 轴上方, 面积有无方向就不必要了. 但如果曲线的一部分在坐标轴下方呢? 例如, 假设 $f(x)=-x^2-2x+3$, 我们所要求的是在 $x=0$ 和 $x=2$ 之间的面积. 因为 $f(0)=3$ 和 $f(2)=-5$, 所以该函数图像如图 16-13 所示.

如果考虑面积的方向, 那么标记为 II 的阴影部分面积就为负的, 这时有

$$\text{有向面积}=\int_0^2(-x^2-2x+3)\mathrm{d}x$$
$$=-\int_0^2 x^2\mathrm{d}x-2\int_0^2 x\mathrm{d}x+3\int_0^2 1\mathrm{d}x.$$

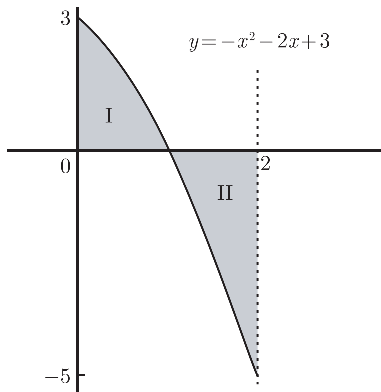


图　16-13

这里利用从前一节学的知识把这个积分分开来写. 因为已知这三个积分的值, 所以有

有向阴影面积

$$=-\frac{8}{3}-2\times(2)+3\times(2)=-\frac{2}{3}\ \text{平方单位}$$

很显然这不是通常的面积, 因为我们所求得的面积是负的! 那么, 怎样求通常的面积呢? 方法是把积分表达式分成几部分, 把在 x 轴下方的面积挑出来, 然后取它们的绝对值. 在上述例子中, 我们需要知道这条曲线与 x 轴的交点. 所以通过解方程 $-x^2-2x+3=0$, 会得到 $x=1$ 或 $x=-3$. 显然, $x=1$ 是我们想要的, 因为它在 0 和 2 之间, 而 -3 不是.

现在, 我们把积分表达式写为两部分：

$$\int_0^1(-x^2-2x+3)\mathrm{d}x\quad\text{和}\quad\int_1^2(-x^2-2x+3)\mathrm{d}x.$$

这分别表示了刚才图像中的两个有向面积 I 和 II. 为计算这两个积分, 我们需要用到本章前面的一些公式：

$$\int_0^1 x^2\mathrm{d}x = \frac{1}{3};\quad \int_0^1 x\mathrm{d}x = \frac{1}{2};\quad \int_0^1 1\mathrm{d}x = 1;$$

$$\int_1^2 x^2\mathrm{d}x = \frac{7}{3};\quad \int_1^2 x\mathrm{d}x = \frac{3}{2};\quad \int_1^2 1\mathrm{d}x = 1.$$

下面的结果留给你去计算.

$$\int_0^1 (-x^2-2x+3)\mathrm{d}x = \frac{5}{3} \quad 和 \quad \int_1^2 (-x^2-2x+3)\mathrm{d}x = -\frac{7}{3}.$$

像我们预测的那样, 第一个积分是正的, 因为面积 I 是在 x 轴上方; 第二个积分是负的, 因为面积 II 是在 x 轴下方. 这两个积分的和是 $-2/3$(平方单位), 这是有向面积. 现在有一个关键点：忽略前面的减号可以求出面积 II 的实际值! 这个方法很管用, 因为这块面积完全在坐标轴的下方. 所以面积 II 的实际面积是 7/3 平方单位, 而面积 I 的面积是 5/3 平方单位, 总面积为 5/3+7/3=4 平方单位. 最有效的方式是, 我们可以取 5/3 和 $-7/3$ 的绝对值, 然后直接相加.

附带着, 我们实际已经证明了

$$\int_0^2 |-x^2-2x+3|\mathrm{d}x = 4.$$

所以, 让我们研究一下为什么带有绝对值的积分求出的就是通常的面积, 就像 $y = |-x^2-2x+3|$ 的图像显示的那样 (如图 16-14 所示).

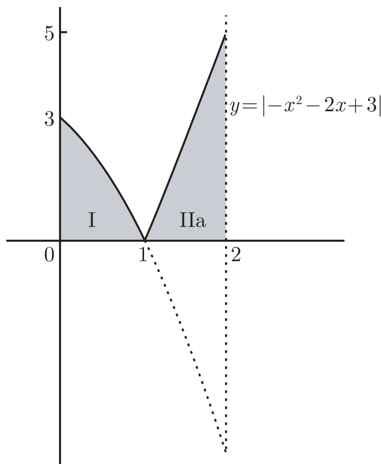


图 16-14

标记为 IIa 的面积与刚才标记为 II 的坐标轴下方的面积关于 x 轴对称, 所以它们面积相同. 该阴影部分的总面积同刚才图像阴影部分的面积一样大.

现在, 我们总结如何求 $y = f(x)$、x 轴和 $x = a$ 及 $x = b$ 所围成的面积. 这个方法同样也适用于下列两个积分, 因为它们都等价于它们所对应的面积.

$$\int_a^b |f(x)|\mathrm{d}x \quad 或 \quad \int_a^b |y|\mathrm{d}x.$$



方法如下所示.

- 找出在 $[a, b]$ 区间内满足函数值为 0 的所有 x 的值.

- 接下来写出以 $f(x)$(而不是 $|f(x)|$) 为被积函数的积分表达式. 第一个积分以 a 开始, 然后以使函数为 0 的最小 x 值结束. 第二个积分以使函数为 0 的最小 x 值开始, 以下一个使函数为 0 的 x 值结束. 以此类推, 直到取遍所有使函数为 0 的 x 值. 最后的积分是以使函数为 0 的最大 x 值开始, 以 b 值结束.
- 分别计算每一个积分.
- 把刚才计算出的每一个积分分别取绝对值, 再把这些数加到一起, 这样就得到了所求的面积.

在 17.6.3 节中, 我们会看到另一个例子. 现在, 我们可以用刚才的方法求物体运动的路程了, 注意是路程不是位移. 实际上, 我们在 16.1 节中得到了下面的式子：

$$\text{路程} = \int_a^b |v(t)|\mathrm{d}t,$$

就像刚才方法中陈述的那样, 我们应用了绝对值.

16.4.2　求解两条曲线之间的面积

假设有两条曲线, 一条在另一条之上, 你想要求它们与 $x = a$ 和 $x = b$ 所围成的面积. 如果曲线是 $y = f(x)$ 和 $y = g(x)$, 前者在后者之上, 图像如图 16-15 所示.

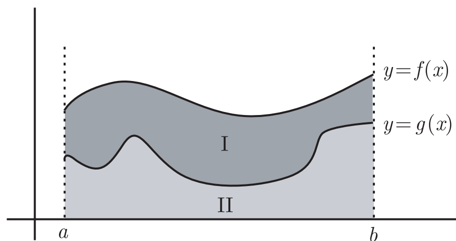


图　16-15

我们要求的面积是标记为 I 的那块面积. 另一方面, 标记为 II 的面积是函数 $y = g(x)$ 与 x 轴所围成的面积, 所以它的面积为

$$\int_a^b g(x)\mathrm{d}x.$$

那么

$$\int_a^b f(x)\mathrm{d}x$$

又是什么呢？

它是上面那个函数与 x 轴所围成的面积, 所以它实际上是两部分的面积和. 所

以我们有

$$\int_a^b f(x)\mathrm{d}x = \int_a^b g(x)\mathrm{d}x + \text{有向面积I}$$

我们可以重写前面的积分表达式, 把这两个积分放到一起, 有

$$\text{有向面积I} = \int_a^b (f(x) - g(x))\mathrm{d}x.$$

所以这两条曲线之间的面积是上边曲线的积分减下边曲线的积分. 例如. 求图 16-16 所示阴影的面积.



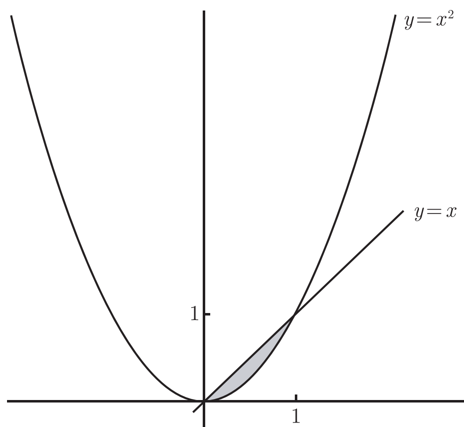


图 16-16

上面的阴影是在 $y = x$ 和 $y = x^2$ 之间所围成的面积. 交点为 x =0 和 x =1, 所以有

$$\text{阴影面积} = \int_0^1 (x - x^2)\mathrm{d}x = \int_0^1 x\mathrm{d}x - \int_0^1 x^2\mathrm{d}x = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}\text{平方单位}.$$



如果区间是从 0 到 2, 结果又是怎样的呢? 图像如图 16-17 所示. 如果把面积表达为

$$\int_0^2 (x - x^2)\mathrm{d}x,$$

那就是大错特错了.

如果计算这个积分, 你会发现它的结果为 $-2/3$, 但这不可能是一个面积的值. 那么问题出现在哪儿呢? 实际上, 仅仅当 x 在区间 0 和 1 之间时, $y = x$ 才在 $y = x^2$ 的上边. 在 $x = 1$ 的右边时, 曲线 $y = x^2$ 是在上边的. 很显然 $x - x^2$ 是不对的, 应该用 $|x - x^2|$ 来替代. 用这种方式, 我们会很确定所求的是实际面积, 无论哪个曲线在上边. 所以我们可以用前面的方法去计算

$$\int_0^2 |x - x^2|\mathrm{d}x.$$

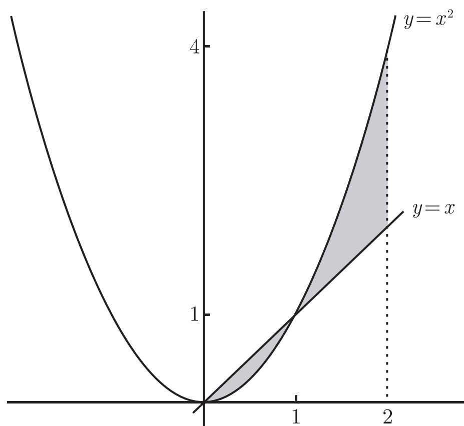


图　16-17

这个不是问题. 首先注意当 $x=0$ 或 $x=1$ 时 $x-x^2=0$, 所以我们考虑下面这两个积分：

$$\int_0^1 (x-x^2)\mathrm{d}x \quad 和 \quad \int_1^2 (x-x^2)\mathrm{d}x.$$

前面积分的结果是 1/6, 但第二个积分的结果是 3/2−7/3= −5/6. 第二个积分是负的, 这个结果是有意义的. 因为当 x 在区间 [1,2] 时, $y=x$ 不在 $y=x^2$ 的上边. 不要管这些, 我们需要做的是把这两个数的绝对值加到一起去：

$$\int_0^2 |x-x^2|\mathrm{d}x = \left|\frac{1}{6}\right| + \left|-\frac{5}{6}\right| = \frac{1}{6}+\frac{5}{6} = 1.$$

所以, 要求的面积是 1 平方单位.

总的来说, 由 $y=f(x)$、$y=g(x)$、$x=a$ 及 $x=b$ 所围成的面积由如下公式给出：

$$\boxed{在函数\ f\ 和\ g\ 之间的面积\ (平方单位)=\int_a^b |f(x)-g(x)|\mathrm{d}x.}$$

如果在区间 $[a, b]$ 内 $f(x)$ 一直是大于或等于 $g(x)$, 那么这个绝对值符号就没有必要了. 否则, 我们可以用 16.4.1 节中的方法去解决这个绝对值问题. 在 17.6.3 节中, 我们将要看到应用这个方法的另一个例子.

16.4.3　求曲线与 y 轴所围成的面积



让我们求这个面积：该面积由 $y=\sqrt{x}$、y 轴以及直线 $y=2$ 围成. 图 16-18 是该面积的图示.

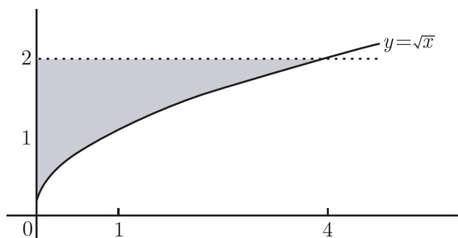


图 16-18

如果我们把面积写成

$$\int_0^2 \sqrt{x}\mathrm{d}x \quad \text{甚至是} \quad \int_0^4 \sqrt{x}\mathrm{d}x$$

将是一个严重的错误.

这两个积分表示的是与 x 轴, 而不是与 y 轴围成的面积. 事实上, 它们分别等价于图 16-19 所示图像的面积.

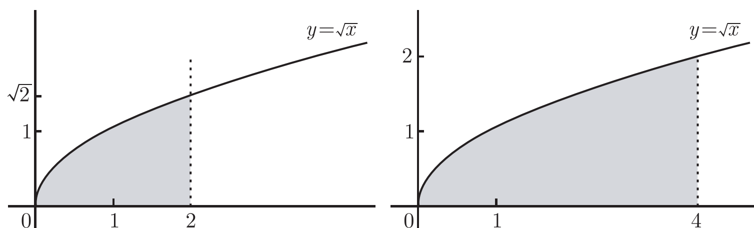


图 16-19

第二个图像更好些, 因为当 $x=4$ 时所对应的 y 值为 2. 但是, 这两个积分表达式都不能正确表达这个面积. 要正确地计算面积, 最好的方法是对 y 求积分, 而不是对 x 求积分. 我们可以把该面积按水平的方向切成条状, 而不是竖条了. 图 16-20 是该图像的例子.

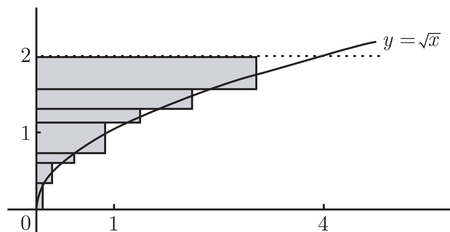


图 16-20

我们以其中一个横条为例, 它的长为 x, 宽为 $\mathrm{d}y$, 如图 16-21 所示.

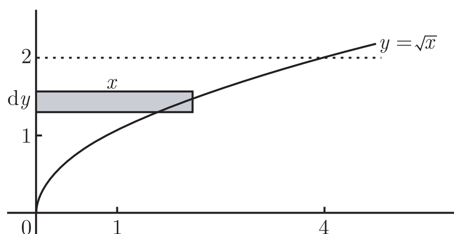


图 16-21

这个小横条的面积为 $x\mathrm{d}y$ 平方单位, 通过积分的方法可以求出整块面积. 在我们的例子中, y 是从 0 到 2(不是到 4), 所要求的面积是 (平方单位)

$$\int_0^2 x\mathrm{d}y.$$

$y=\sqrt{x}$, 可知 $x=y^2$, 所以上述积分表达式可写为

$$\int_0^2 y^2\mathrm{d}y.$$

这同我们以前的积分表达式

$$\int_0^2 x^2\mathrm{d}x$$

没有什么不同, 只是虚拟变量由 x 变成了 y. 这个改变对我们的积分结果并没有影响, 该积分依然为 8/3, 所以这块面积为 8/3 平方单位. 要想弄清楚这一点, 让我们重新看看刚才的面积. 可以发现需要做的是把该图像依 $y=x$ 这条直线对称翻转, 这样得到函数 $y=x^2$ 从 $x=$0 到 $x=$2 的面积. 我们所需要做的仅仅是把 x 和 y 对换. 当然, 如果 $y=f(x)$, 反函数是存在的, 那么我们有 $x=f^{-1}(y)$, 所以我们可以把上述观点总结如下.

> 如果 f 存在反函数, $\displaystyle\int_A^B f^{-1}(y)\mathrm{d}y$ 就是由函数 $y=f(x)$、直线 $y=A$ 和 $y=B$ 以及 y 轴所围成的面积 (平方单位).

如果你喜欢, 可把上述积分写为

$$\int_A^B x\mathrm{d}y,$$

这是因为当 $y=f(x)$ 时 $x=f^{-1}(y)$. 而且, 请注意积分的上下限, 我用的是大写字母 A 和 B—— 这样做的目的是强调我们是对 y 求积分, 而不是对 x 求积分. 所以刚才的例子中, 积分上下限是从 0 到 2 而不是从 0 到 4. 因为 $f(x)=\sqrt{x}$, 我们可以说 $f^{-1}(x)=x^2$. 所以上述公式也可以改写为

$$\int_A^B f^{-1}(y)\mathrm{d}y = \int_0^2 y^2\mathrm{d}y,$$

结果为 8/3, 就像我们刚才算的那样.

16.5 估算积分

这有个非常简单但很实用的原则: **当一个函数一直都大于另一个函数时, 它的积分也一直大于另一个函数的积分**. 让我们看看图 16-22 所示的图像.

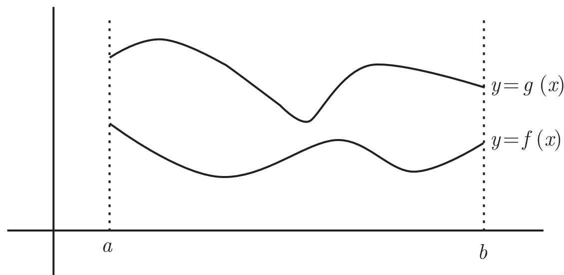


图 16-22

在区间 $[a, b]$ 内, 函数 g 一直都在函数 f 的上方. (我们在 16.4.1 节中见过的这两个函数, 情况正好相反!) 在任何情况下, 函数 $y = f(x)$ 与 x 轴所围成的面积都要比函数 $y = g(x)$ 与 x 轴所围成的面积小. 用符号可表示如下.

> 如果对于在区间 $[a, b]$ 内的所有 x 都有 $f(x) \leqslant g(x)$, 那么就有
> $$\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x.$$

即使这两个曲线都在 x 轴的下方, 这个结论也是成立的, 因为我们使用的是有向面积. 例如, 如果函数 f 是在 x 轴的下方, 函数 g 是在 x 轴上方, 这时积分 $\int_a^b f(x)\mathrm{d}x$ 为负, 而 $\int_a^b g(x)\mathrm{d}x$ 为正, 所以上述不等式依然是成立的.



如果使用黎曼和, 那么证明上述结论是非常容易的. 不用考虑细节, 我们仅仅需要考虑划分, 并注意到对于任意一个 j 都有 $f(c_j) \leqslant g(c_j)$, 所以函数 f 的黎曼和小于函数 g 的黎曼和. 我把证明过程留给你了.

我们可以用速度和位移更好地解释上述公式. 假设在同一地点有两辆车同时出发. 第一辆车在时刻 t 的速度为 $f(t)$, 第二辆车在时刻 t 的速度为 $g(t)$. 因为速度的积分是位移, 所以上述结论中的公式可以解释为, 如果第一辆车的速度一直比第二辆车的速度小, 那么可以说第一辆车的位移要比第二辆车的位移小. 你若这样考虑, 就很容易理解了. 如果我们以右方向为正方向, 那么第一辆车永远在第二辆车的左边, 它永远不可能到达第二辆车的右边.

一个简单的估算

使用上述不等式, 我们不用计算定积分的值也能估算一个定积分有多大或多小. 例如, 我们要估计 $\int_a^b f(x)\mathrm{d}x$ 的值, 也就是图 16-23 所示的面积.

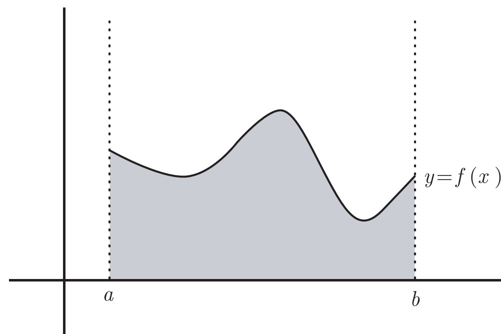


图　16-23

设 M 为函数 $f(x)$ 在 $[a, b]$ 区间的最大值, 设 m 为其在该区间的最小值. 我们分别把 $y = M$ 和 $y = m$ 两条直线画出来, 图像如图 16-24 所示.

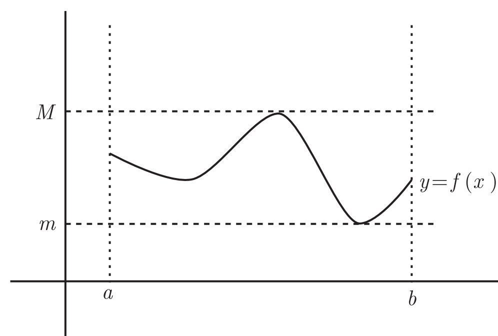


图　16-24

注意, 我们要计算的面积是在直线 $y = M$ 和 $y = m$ 之间. 这点通过绘制更多的函数图像很容易看出来, 如图 16-25 所示.

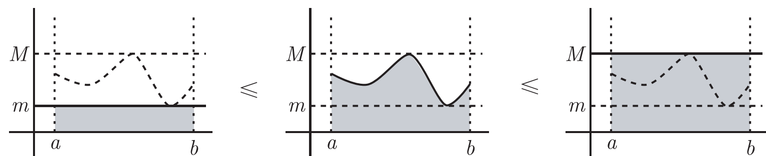


图　16-25

我们可以很容易地求出图 16-25 中左图和右图中长方形的面积. 对于左边的长方形, 底为 $(b - a)$, 高为 m, 所以它的面积为 $m(b - a)$ 平方单位. 对于右边的长方

形, 底仍然为 $(b-a)$, 但它的高为 M, 所以面积为 $M(b-a)$ 平方单位. 由上述图像得到以下结论.

> 如果对于在 $[a, b]$ 区间内的所有 x 有 $m \leqslant f(x) \leqslant M$, 那么
> $$m(b-a) \leqslant \int_a^b f(x)\mathrm{d}x \leqslant M(b-a).$$

当然, 这里我们两次应用了上一节的原则. 下面来看一个使用这个结论的例子. 假设我们要想知道积分

$$\int_0^{1/2} \mathrm{e}^{-x^2}\mathrm{d}x$$

的值是多少. $y=\mathrm{e}^{-x^2}$ 的函数图像是非常有名的钟形曲线, 它处处可见, 特别是在概率论和统计学中. 我们计算图 16-26 中阴影部分的面积.

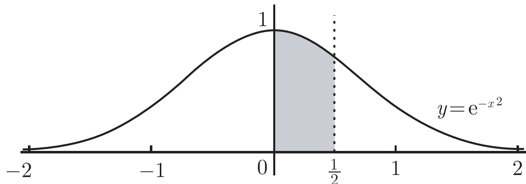


图 16-26

即使用上接下来三章中所有计算积分的方法, 我们都不能计算出该积分的准确值. 事实上, 如果不用积分符号或求和, 我们再也找不到更好的方法去表达这个积分值了. 但至少我们可以用刚才的原理估算一下它的值.

我们需要找到在 $\left[0, \dfrac{1}{2}\right]$ 区间内函数 $y=\mathrm{e}^{-x^2}$ 的最大值和最小值. 通过链式求导法则, 我们有 $\mathrm{d}y/\mathrm{d}x=-2x\mathrm{e}^{-x^2}$, 当 x 为 0 时导数为 0, 其余情况为负值. 这样可以说 $y=\mathrm{e}^{-x^2}$ 在 $\left[0, \dfrac{1}{2}\right]$ 区间为减函数, 所以最大值出现在 $x=0$ 处, 最小值出现在 $x=1/2$ 处. 把这些值代入, 可以得到最大值为 $\mathrm{e}^{-0^2}=1$, 最小值为 $\mathrm{e}^{-(1/2)^2}=\mathrm{e}^{-1/4}$. 也就是说, 在区间 $\left[0, \dfrac{1}{2}\right]$ 中有

$$\mathrm{e}^{-1/4} \leqslant \mathrm{e}^{-x^2} \leqslant 1.$$

根据上面的原理, 把 $a=0, b=\dfrac{1}{2}$ 代入可得

$$\mathrm{e}^{-1/4} \times \left(\frac{1}{2}-0\right) \leqslant \int_0^{1/2} \mathrm{e}^{-x^2}\mathrm{d}x \leqslant 1 \times \left(\frac{1}{2}-0\right).$$

所以, 可以说要求的积分值在 $\dfrac{1}{2}\mathrm{e}^{-1/4}$ 和 $\dfrac{1}{2}$ 之间. 通过图 16-27, 我们可以更清楚地

看到, 一个值估算过高, 一个值估算过低.

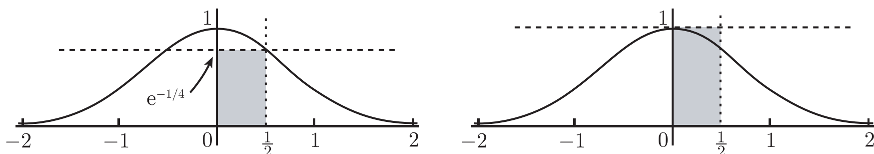


图　16-27

这两个长方形的面积分别为 $\frac{1}{2}e^{-1/4}$ 和 $\frac{1}{2}$ 平方单位.

上面的估算很不精确. 我们可以使用更多的长方形做更精确的估算, 或者使用梯形、抛物线一样的小竖条等奇特形状. 更多信息参见附录 B.

16.6　积分的平均值和中值定理

最后我们讨论平均速度的问题. 是的, 在单位时间内, 我们可以说速率的值等于路程, 也可以说速度的值等于位移. 但这段陈述成立的前提是速度为常数; 否则, 就像在 5.2.3 节讲述的那样, 需要引入**平均速度**.

我们已经了解, 使用微分可以在已知某时间段位移的前提下求即时速度. 使用积分, 可以在已知某时间段即时速度的前提下求位移. 当然, 在已知某时间段即时速度的前提下, 也可以求出平均速度. 你所需要做的是求出位移, 然后用这个值除以总时间. 如果时间是从 a 到 b, 在时刻 t 的速度是 $v(t)$, 那么我们有

$$\text{位移} = \int_a^b v(t)\mathrm{d}t.$$

因为总的时间为 $b-a$, 所以有

$$\text{平均速度} = \frac{\text{位移}}{\text{总时间}} = \frac{1}{b-a}\int_a^b v(t)\mathrm{d}t.$$

总的来说, 我们可以在区间 $[a, b]$ 内, 将可积函数 f 的平均值定义为：

$$\boxed{\text{函数 } f \text{ 在区间 } [a,b] \text{ 内的平均值} = \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x.}$$

例如, 求函数 $f(x)=x^2$ 在 $[0, 2]$ 区间上的平均值是多少? 很简单,

$$\text{平均值} = \frac{1}{2-0}\int_0^2 x^2\mathrm{d}x = \frac{1}{2}\times\frac{8}{3} = \frac{4}{3}.$$

所有要做的是, 用积分结果除以积分上下限的差.

来看看这个定义的几何解释. 我们把函数 f 在区间 $[a,b]$ 的平均值记为 f_{av}, 图 16-28 是关于 $y=f(x)$ 和 $y=f_{\mathrm{av}}$ 图像的例子.

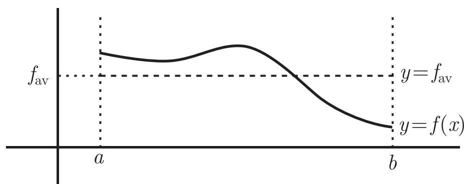


图 16-28

注意, 如果 f_{av} 仅仅为一常数, 则 $y = f_{av}$ 的图像是一条水平线. 现在, 使用上述公式, 我们有

$$f_{av} = \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x.$$

两边同时乘以 $(b-a)$, 可得

$$\int_a^b f(x)\mathrm{d}x = f_{av} \times (b-a).$$

这实际上是在说, 图 16-29 中两个图像阴影部分的面积是相等的,

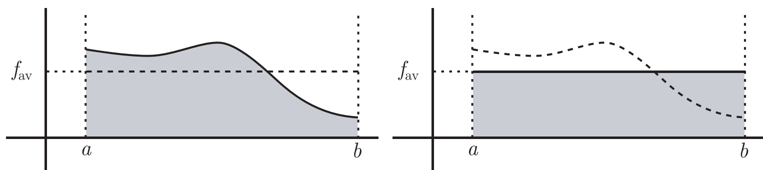


图 16-29

由图 16-29 可见, 右侧图像中长方形的高为 f_{av} 单位, 底为 $(b-a)$ 单位, 所以它的面积为 $f_{av} \times (b-a)$ 平方单位. 你可以这样考虑它：假设你拨动鱼缸里的水, 水面在某一时刻就像函数 $y=f(x)$, 等水平静下来, 其表面就像$y=f_{av}$ 这条水平线.

积分的中值定理

在上述图像中, 水平线 $y = f_{av}$ 与函数 $y = f(x)$ 有交点, 我们将其横坐标记为 c, 如图 16-30 所示.

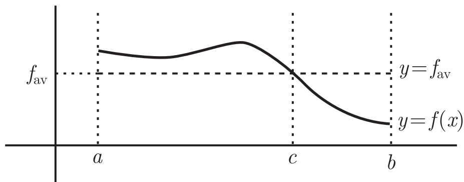


图 16-30

所以我们有 $f(c) = f_{av}$. 可以得出这样的结论：如果函数 f 是连续的, 那么总会有这样一个数 c.

> **积分的中值定理：**如果函数 f 在闭区间 $[a,b]$ 上连续, 那么在开区间 (a,b) 内总有一点 c, 满足 $f(c)=\dfrac{1}{b-a}\displaystyle\int_a^b f(x)\mathrm{d}x$.



简言之, 连续函数在一段区间内至少一次达到它的平均值. 例如, 在上一节中我们看到, 函数 $f(x)=x^2$ 在区间 [0, 2] 的平均值为 4/3. 根据上述定理, 我们可以说一定有一个数 c 在区间 [0, 2] 满足 $f(c)=4/3$. 因为 $f(c)=c^2$, 可以知道 $c=\sqrt{4/3}$ 是在区间 [0, 2] 上的一个解 (另一个解 $c=-\sqrt{4/3}$ 不在该区间内).

如果从速度的角度考虑上述定理, 我们可以说在区间 $[a,b]$ 内有一点 c 满足 $v(c)=v_{\mathrm{av}}$. 也就是说, 在任何一段旅途中都有一个时刻 c, 使得这个时刻的速度 $v(c)$ 等于该段路程的平均速度 v_{av}. 无论你怎样努力求证, 在任何一段旅途中, 至少会有这样一个时刻, 其即时速度等于该段路程的平均速度. 至少会有一个这样的时刻, 不可能一个都没有. 假设你在第一小时的速度为 45 英里/小时, 在第二小时的速度为 55 英里/小时, 该段路程的平均速度为 50 英里/小时, 那么在该段时间内一定会有某一个时刻的速度为 50 英里/小时, 这个时刻可能出现在从 45 英里/小时到 55 英里/小时的加速过程中.

为什么上述定理也叫作中值定理呢? 毕竟, 我们已经有了一个中值定理. 如果重新看一下 11.3 节讨论过的定理, 你会发现我们两次得到的是同样的结论：在任何一段旅途中, 都有某一时刻的即时速度等于平均速度. 这两个定理中唯一的不同是：在前一个版本中, 我们是用位移 – 时间图像中的斜率来解释的; 而现在, 我们使用速度 – 时间图像中的面积来解释.



现在来看看这个定理为什么是成立的. 如 16.5 节所述, 我们设 M 为函数在 $[a,b]$ 区间的最大值, m 为函数在 $[a,b]$ 区间的最小值. f_{av} 可能比 M 大吗? 如果它比 M 大, 那么情况将会如图 16-31 所示.

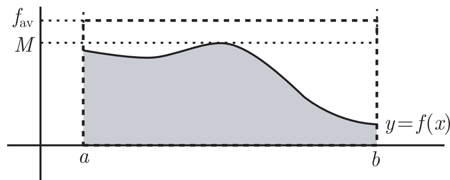


图　16-31

虚线部分所围成的长方形面积不可能等于阴影部分的面积, 因为长方形包含了阴影区域! 所以这种情况不可能出现. 同样, f_{av} 也不可能比最小值 m 小. 它一定会在 m 和 M 之间. 介值定理告诉我们, 函数 f 可取 m 和 M 之间的任意值 (你知道为什么吗), 所以 f 在某一时刻一定等于平均值 f_{av}. 也就是说, 一定有某个数 c

满足 $f(c) = f_{\mathrm{av}}$, 所以该定理是正确的. 在 17.8 节中, 我们将使用该定理证明微积分学的第一基本定理.

16.7 不可积的函数

16.2 节曾提及, 如果函数 f 为有界函数并在区间 $[a, b]$ 上有有限个不连续点, 那么函数 f 是可积的; 也就是说, 定积分 $\int_a^b f(x)\mathrm{d}x$ 存在. 顺便提一下, 不连续是不可导的一种情况; 也就是说, 如果函数在 $x = a$ 点不连续, 那么它在该点也不可导 (参见 5.2.11 节). 积分同可导的情况有所不同, 即使是不连续的函数, 只要它有有限个不连续点也是可积的. 现在, 让我们看一个有太多个不连续点的函数的积分情况.

首先, 我们回忆一下有理数的定义. 有理数可以被写成 p/q 形式, 其中 p 和 q 为整数 (它们没有公约数), 而无理数就不可能写成这种形式. 现在, 对于区间 [0, 1] 内的数 x, 我们设

$$f(x) = \begin{cases} 1 & \text{如果 } x \text{ 是有理数.} \\ 2 & \text{如果 } x \text{ 是无理数.} \end{cases}$$

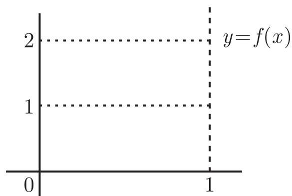


图 16-32

这是一个很奇怪的函数. 在 0 和 1 之间有太多的有理数和无理数. 事实上, 每两个有理数之间都有一个无理数; 每两个无理数之间也有一个有理数! 所以当我们试着绘制函数 $y = f(x)$ 的图像时, 可能会想到如图 16-32 的图像.

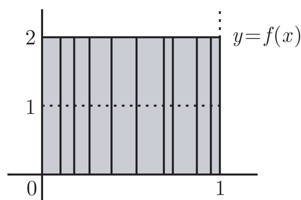


图 16-33

函数 $f(x)$ 的值在高度 1 和 2 之间以超乎你想象的速度来回跳跃. 在 1 和 2 这两条线段间有很多不连续处, 我们说有很多不连续点. 这个函数实际上在任何一点都不连续. 那么积分 $\int_0^1 f(x)\mathrm{d}x$ 究竟为多少呢? 让我们取黎曼上和以及黎曼下和. 在此把区间 [0, 1] 分成许多小区间. 无论这些子区间的宽有多小, 小竖条中都会有一些无理数点. 所以求上和会如图 16-33 所示.

为了求上和, 每一个长方形的高一定要是 2, 即使这个长方形很窄. 注意, 无论其中有多少个长方形, 所有长方形的面积和为 2 平方单位, 因为我们是对一个 1 乘以 2 的长方形进行划分的. 这样就有

$$\lim_{\text{最大区间}\to 0}(\text{取黎曼上和}) = \lim_{\text{最大区间}\to 0} 2 = 2.$$

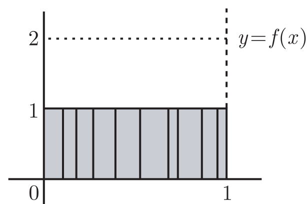


图 16-34

相似地, 对同样的划分求下和, 每个长方形的高为 1 单位. 毕竟, 无论长方形的宽多小, 它的底 (在 x 轴上) 都包含一个有理数. 对于所有的有理数, 该函数的高为 1. 所以求下和如图 16-34 所示.

现在该面积为 1 平方单位, 因为这些小分区填充的大长方形为 1 乘以 1. 所以我们已经证明

$$\lim_{\text{最大区间}\to 0}(\text{取黎曼下和}) = \lim_{\text{最大区间}\to 0} 1 = 1.$$

当最大区间趋于 0 时, 这个极限取黎曼上和和取黎曼下和是不同的. 对于连续函数, 这种情况不会出现. 但对于一些不连续的函数, 这种情况时有发生! 唯一的结论是, 函数在区间 [0, 1] 上不可积. 我们说函数 f 是*不可积的*. 实际上有一种方法可以求这种函数的积分, 叫作**勒贝格积分**(与黎曼积分相对), 它超出了本书的讨论范围. 所以, 我们不用考虑这种不正常的积分, 而是要寻求求解正常、连续函数的定积分的好方法.

第 17 章　微积分基本定理

我们现在来讨论微积分中的关键部分 —— 微积分基本定理, 它不仅提供了一种不用黎曼和就可以求解定积分的方法, 同时也展示了微分和积分的关系. 不多说了, 这一章我们将要学习:

- 用另一个函数的积分形式来表示的函数;
- 第一基本定理, 以及反导数的基本思想;
- 第二基本定理;
- 不定积分和它们的性质.

在介绍完所有这些理论之后, 我们将针对下面的知识点举出若干例子:

- 以第一基本定理为基础的问题;
- 计算不定积分;
- 计算定积分以及使用第二基本定理计算面积.

17.1　用其他函数的积分来表示的函数

在上一章中, 我们使用黎曼和证明了

$$\int_0^1 x^2 \mathrm{d}x = \frac{1}{3} \quad \text{和} \quad \int_0^2 x^2 \mathrm{d}x = \frac{8}{3}.$$

(实际上, 我们仅仅证明了第二个, 第一个留给你了!) 遗憾的是, 黎曼和方法太烦琐了, 最好能找到一个相对简单的方式. 为什么我们在那儿停下来了呢? 让我们试着计算

$$\int_0^{\text{任意数}} x^2 \mathrm{d}x.$$

在此, 我们让极限上限为变量. 最常用的变量是 x, 但是你不能把这个积分写成

$$\int_0^x x^2 \mathrm{d}x,$$

除非你想造成混乱局面. 毕竟, x 是虚拟变量, 实际上不是一个变量. 我们重新开始, 这次使用 t 为虚拟变量. 首先, 我们有

$$\int_0^1 t^2 \mathrm{d}t = \frac{1}{3} \quad \text{和} \quad \int_0^2 t^2 \mathrm{d}t = \frac{8}{3}.$$

请记住, 虚拟变量用什么字母都无所谓 —— 我们已经重新命名 x 轴为 t 轴. 实际

面积并没有改变. 现在我们考虑积分

$$\int_0^x t^2 \mathrm{d}t.$$

如果把 $x=1$ 代入这个积分表达式, 会得到 $\int_0^1 t^2 \mathrm{d}t$, 积分结果为 1/3; 如果把 $x=2$ 代入, 会得到 $\int_0^2 t^2 \mathrm{d}t$, 积分结果为 8/3. 为什么得出这个表达式就不往下计算了呢? 你可以把任意值放到 x 的位置, 得到不同的积分. 上述积分表达式实际上是一个以积分上限 x 为变量的函数. 我们用 F 来标记这个函数, 这样有

$$F(x) = \int_0^x t^2 \mathrm{d}t.$$

可以发现, $F(1) = 1/3, F(2) = 8/3$. 那么 $F(0)$ 是多少呢? 来看下面式子:

$$F(0) = \int_0^0 t^2 \mathrm{d}t.$$

在 16.3 节中我们已经知道, 对于积分上下限都一样的积分表达式, 该积分的结果为 0. 也就是说, 我们知道 $F(0) = 0$. 不走运的是, 求解其他的 F 值并不是很容易, 例如 $F(9)$、$F(-7)$ 或 (1/2). 在下一节中, 我们将要研究这个问题. 与此同时, 怎样用文字来描述 $F(x)$ 呢? 准确地说, 它应该是曲线 $y = t^2$、t 轴与 $t = x$ 所围成的有向面积.

我们可以用两种方式推广这个问题. 首先, 积分下限不一定是 0. 你可以定义另一个函数:

$$G(x) = \int_2^x t^2 \mathrm{d}t.$$

这个积分表达式也可以用面积 (平方单位) 来解释, 它是由曲线 $y = t^2$、t 轴以及两条垂线 $t=2$ 和 $t = x$ 所围成的面积. 那么 $G(2)$ 为多少呢? 来看下式:

$$G(2) = \int_2^2 t^2 \mathrm{d}t = 0,$$

因为积分上下限是一样的. 那么 $G(0)$ 是多少呢? 我们有

$$G(0) = \int_2^0 t^2 \mathrm{d}t.$$

16.3 节曾讲述怎样计算这个积分, 你可以交换积分的上下限, 然后再在积分表达式的前边加上一个负号. 所以有

$$G(0) = \int_2^0 t^2 \mathrm{d}t = -\int_0^2 t^2 \mathrm{d}t = -\frac{8}{3}.$$

事实上, 在函数 F 和 G 之间有一种奇妙的关系. 首先, 这两个函数是:

$$F(x) = \int_0^x t^2 \mathrm{d}t \quad 和 \quad G(x) = \int_2^x t^2 \mathrm{d}t.$$

我们从 $t=2$ 这点分解第一个积分表达式, 可参见 16.3 节. 可得

$$\int_0^x t^2\mathrm{d}t=\int_0^2 t^2\mathrm{d}t+\int_2^x t^2\mathrm{d}t.$$

左边是 $F(x)$. 右边的第一项是 8/3, 第二项是 $G(x)$. 这样, 就证明了

$$F(x)=\frac{8}{3}+G(x).$$

也就是说, F 和 G 的差是 8/3. 我们可以做得更多. 假设 a 是任意固定的数, 设

$$H(x)=\int_a^x t^2\mathrm{d}t.$$

如果从 $t=a$ 而不是 $t=2$ 分解函数 F, 会得到

$$F(x)=\int_0^x t^2\mathrm{d}t=\int_0^a t^2\mathrm{d}t+\int_a^x t^2\mathrm{d}t.$$

右侧的第二项恰恰就是 $H(x)$, 所以我们已经证明了

$$F(x)=\int_0^a t^2\mathrm{d}t+H(x).$$

这是什么呢? 实际上, $\int_0^a t^2\mathrm{d}t$ 是个常数 —— 它不因 x 的变化而变化! 尽管我们没有确定 a 的值, 但说过 a 是一个常数, 所以这个积分结果一定是个常数. 这样就证明了

$$F(x)=H(x)+C,$$

其中 C 是一个常数, 由 a 而不是 x 决定. 这个方法的基本思想是把积分下限从一个常数换至另一个常数, 这对整个表达式没有太大的影响.

我们的第二个解释是被积函数不一定是 t^2, 它可以是关于 t 的任意连续函数. 假设被积函数是 $f(t)$, 如果 a 是任意常数, 我们定义

$$F(x)=\int_a^x f(t)\mathrm{d}t.$$

例如, 如果 $a=0, f(t)=t^2$, 则可以从上述定义得到原始函数 F. 总的来说, 对任何数 x, 函数 $F(x)$ 的值都是一个有向面积 (平方单位), 该区域是由曲线 $y=f(t)$、t 轴以及 $t=a$ 和 $t=x$ 两条垂线所围成的. 图 17-1 是关于不同 x 的 3 种情况图示.

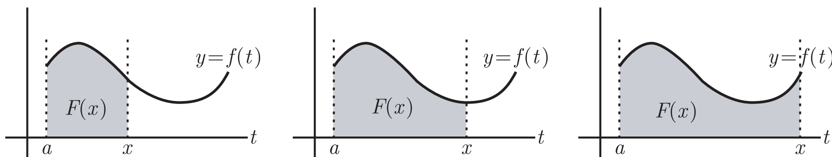


图 17-1

上述图像让人想到了窗帘, 左边固定, 右边移动. 不真实的一面是, 窗帘杆高低不平, 除非 f 是个常函数! 在任何情况下, 请注意函数 F 主要是由被积函数 $f(t)$ 和常数 a 决定的. 通过刚才的分割法可知, 改变 a 的值仅仅使函数值增加或减少一个常数, 并没有太大的影响. 后面几节将会体现出所有这些思想的重要性.

17.2　微积分的第一基本定理

我们的目的是不用黎曼和来求积分

$$\int_a^b f(x)\mathrm{d}x.$$

我们要做 3 件并不显而易见的事情.

(1) 首先, 把虚拟变量改为 t, 把上述积分表达式写为 $\int_a^b f(t)\mathrm{d}t$. 像上一节那样, 没什么不同 —— 用什么来表示虚拟变量无关紧要.

(2) 现在, 用变量 x 来替代 b 从而得到一个新的函数 F, 定义 $F(x) = \int_a^x f(t)\mathrm{d}t$. 这就是我们在上一节见过的函数. 最终要求函数 $F(b)$ 的值, 即第 (1) 步中的积分. 但是, 我们首先来看看该怎样理解函数 F.

(3) 现在有了这个新函数 F, 它像是我们刚刚得到的一个新玩具. 在之前章节中, 我们已经花了很多时间求解函数的导数, 这次将对 x 求这个函数的导数. 考虑

$$F'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\mathrm{d}t.$$

理解 $F'(x)$ 的实质将会帮助我们求解 $F(x)$. 一旦找到这个答案, 就能计算出 $F(b)$, 这就是我们要求解的积分.

表达式

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\mathrm{d}t.$$

看起来可能很奇怪, 让我们看看怎样才能拆开它. 选你最喜欢的变量 x 并求解 $F(x)$. 这时微微变换一下 x—— 把它变为 $x+h$, 其中 h 是个很小的数. 所以, 现在的函数值是 $F(x+h)$. 这种情况的图像如图 17-2 所示.

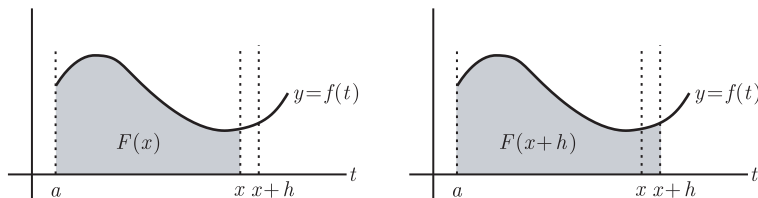


图　17-2

可以看到, x 和 $x+h$ 非常接近, 它们所对应的函数值 $F(x)$ 和 $F(x+h)$ 也非常接近 —— 它们分别表示上图中阴影部分的面积. 现在对 F 求导, 我们有

$$\lim_{h\to 0}\frac{F(x+h)-F(x)}{h}.$$

$F(x+h)-F(x)$ 的差就是图 17-2 中两阴影部分面积的差, 也就是那个小竖条的阴影面积 (顶部是弯曲的), 该面积在 $t=x$ 和 $t=x+h$ 之间, 如图 17-3 所示.

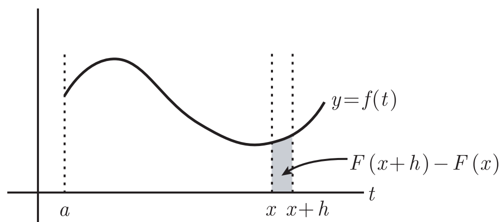


图 17-3

我们可以通过从 $t=x$ 处分解这个积分来计算函数 $F(x+h)$ 的值, 像这样:

$$F(x+h)=\int_a^{x+h}f(t)\mathrm{d}t=\int_a^{x}f(t)\mathrm{d}t+\int_x^{x+h}f(t)\mathrm{d}t=F(x)+\int_x^{x+h}f(t)\mathrm{d}t.$$

通过整理可得

$$F(x+h)-F(x)=\int_x^{x+h}f(t)\mathrm{d}t,$$

这就是小竖条的阴影部分面积 (平方单位). 实际上, 这并不是一个竖条, 因为它的顶是弯曲的. 但当 h 很小的时候, 它几乎就是个竖条了. 该竖条左边的高度为 $f(x)$ 单位, 所以可以用计算长方形面积的方法来估算该竖条的面积, 它的底从 x 到 $x+h$, 高从 0 到 $f(x)$, 如图 17-4 所示.

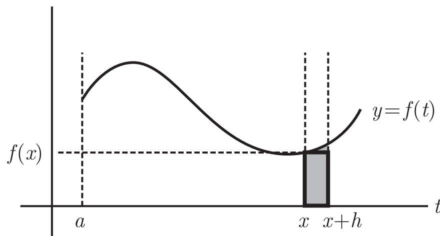


图 17-4

这样该长方形的底为 h 单位, 高为 $f(x)$ 单位, 所以面积是 $hf(x)$ 平方单位. 如果 h 很小, 那么这就是对这个积分的一个非常好的估算. 也就是

$$F(x+h)-F(x)=\int_x^{x+h}f(t)\mathrm{d}t\approx hf(x).$$

两边同时除以 h , 得到

$$\frac{F(x+h)-F(x)}{h} \approx f(x).$$

当 h 非常接近于 0 时, 这个估算就会很准确. 也就是说, 当 h 趋于 0 时, 这个估算是精确的：

$$\lim_{h\to 0}\frac{F(x+h)-F(x)}{h} = f(x).$$

我们会在 17.8 节中看到, 上述公式是正确的. 我们可以总结为

$$F'(x) = f(x).$$

总结如下.

微积分的第一基本定理：如果函数 f 在闭区间 $[a,b]$ 上是连续的, 定义 F 为

$$F(x) = \int_a^x f(t)\mathrm{d}t,\ x \in [a,b]$$

则 F 在开区间 (a,b) 内是可导函数, 而且 $F'(x) = f(x)$.

简而言之, 可以总结为

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\mathrm{d}t = f(x).}$$

我们把这个奇怪的表达式化简为 $f(x)$!

关于这个表达式要关注的一点是, a 出现在积分下限而不是积分上限. 这一点确实很有用, 信不信由你. 假设 A 是区间 $(a,\, b)$ 中的某个数, 并且

$$F(x) = \int_a^x f(t)\mathrm{d}t \quad 和 \quad H(x) = \int_A^x f(t)\mathrm{d}t.$$

如我们在 17.2 节见过的那样, F 和 H 的差是个常数 C：

$$F(x) = H(x) + C$$

如果对两边分别求导, 这个常数就消失了, 会得到 $F'(x) = H'(x)$ (x 在区间 $(a,\, b)$ 内). 所以, 常数 a 的选择不会影响这个求导的结果. 拿窗帘来做比喻, 我们需要考虑的是拉窗帘的速度有多快, 以及右侧拉点的位置放多高. 而左侧的固定点并不影响整体的拉动效果.

反导数的引入

现在稍事休息. 我们以一些变量为 t 的函数以及常数 a 开始, 然后建立了一个以 x 为变量的新函数 F. 对 F 求导, 可以得到原来的函数 f, 但现在我们要以 x 为变量而不是 t 来计算它. 很奇怪吧!



是的, 很奇怪, 但很有用. 它实际上解决了我们的一个大问题. 让我们看看它是怎样解决的. 假设 $f(t) = t^2, a = 0$, 所以有

$$F(x)=\int_0^x t^2\mathrm{d}t.$$

微积分的第一基本定理告诉我们 $F'(x)=f(x)$. 因为 $f(t)=t^2$, 所以有 $f(x)=x^2$; 也就是说, $F'(x)=x^2$. 换一种说法, 函数 F 的导数为 x^2. 我们说 F 是 x^2 的反导数 (关于 x). 你能想到其他的函数, 它的导数为 x^2 吗? 这里有一些:

$$G(x)=\frac{x^3}{3},\quad H(x)=\frac{x^3}{3}+7,\quad J(x)=\frac{x^3}{3}-2\pi.$$

在每一种情况下, 我们都可以发现导数为 x^2. 事实上, 任何形式为 $\frac{x^3}{3}+C$(其中 C 为任意常数) 的关于 x 的函数都是 x^2 的反导数. 还有其他的吗? 答案是否定的! 我们在 11.3.1 节已经得到了这个结论. 如果两个函数有相同的导数, 那么它们的差是个常数. 这就是说, 所有反导数之间的差都是一个常数. 因为其中的一个反导数是 $x^3/3$, 所以任何其他反导数一定是 $x^3/3+C$, C 是任意常数. 等一下, 刚才那个奇怪的函数也是 x^2 的反导数. 也就是说对于某个常数 C 有

$$F(x)=\int_0^x t^2\mathrm{d}t=\frac{x^3}{3}+C.$$

现在我们所要做的是找到 C. 我们知道

$$F(0)=\int_0^0 t^2\mathrm{d}t=0.$$

所以有

$$0=\frac{0^3}{3}+C.$$

这就是说 C =0. 现在我们找到了一直在寻找的公式:

$$\int_0^x t^2\mathrm{d}t=\frac{x^3}{3}.$$

最后, 要从 0 到任意数对 t^2 求积分. 具体情况是, 如果用 1 和 2 分别替代 x, 就会得到熟悉的公式:

$$\int_0^1 t^2\mathrm{d}t=\frac{1^3}{3}=\frac{1}{3}\quad 和\quad \int_0^2 t^2\mathrm{d}t=\frac{2^3}{3}=\frac{8}{3}.$$

这其实可以更简单, 我们将在下一节介绍. 首先, 我将要介绍另一个重要的知识点. 我们现在用一种方式去建立任何一个连续函数的反导数. 例如, e^{-x^2} 的反导数为多少呢? 我们把变量 x 换为 t, 选一个你喜欢的数作为积分下限 (我们暂时选 0), 求积分得到 e^{-x^2} 的反导数为



$$F(x)=\int_0^x \mathrm{e}^{-t^2}\mathrm{d}t.$$

常数 0 可以被任意其他数替代, 替代后的式子依然成立. 当然, 对于每一个不同的积分下限, 你会得到一个不同的反导数.

17.3　微积分的第二基本定理

上一节 $f(t)=t^2$ 的例子告诉了我们怎样求解 $\int_a^b f(t)\mathrm{d}t$. 首先, 我们知道被定义为

$$F(x)=\int_a^x f(t)\mathrm{d}t$$

的函数 F 是函数 f (关于 x) 的反导数. 我们真地很想求解 $F(b)$, 因为

$$F(b)=\int_a^b f(t)\mathrm{d}t.$$

我们知道

$$F(a)=\int_a^a f(t)\mathrm{d}t=0,$$

因为积分上下限是一样的.

现在, 假设对于函数 f 有其他的反导数, 我们称之为 G. 这时 F 和 G 之间的唯一不同是相差一个常数, 所以有 $G(x)=F(x)+C$. 如果用 a 替代 x, 就有 $G(a)=F(a)+C$; 因为由上述计算知 $F(a)$=0, 所以有 $G(a)=C$. 这就是说

$$F(x)=G(x)-C=G(x)-G(a).$$

如果用 b 替代 x, 我们有

$$F(b)=G(b)-G(a).$$

换一种方式表达, 即

$$\int_a^b f(t)\mathrm{d}t=G(b)-G(a).$$

这对于任何反导数 G 都是成立的. 注意, 表达式里已经没有 x 了. 现在要做的是把虚拟变量变回 x, 再把函数字母由 G 变回 F, 这样就有了如下的结论.

微积分的第二基本定理: 如果函数 f 在闭区间 $[a,b]$ 上是连续的, F 是 f 的任意一个反导数 (关于 x), 那么有

$$\int_a^b f(x)\mathrm{d}x=F(b)-F(a).$$

在实践中, 我们通常把等式右边写成 $F(x)\Big|_a^b$ 的形式. 也就是说, 设

$$F(x)\Big|_a^b=F(b)-F(a).$$



以计算

$$\int_1^2 x^2\mathrm{d}x$$

为例, 第一步我们寻找 x^2 的一个反导数. 我们已经知道 $x^3/3$ 是一个反导数, 所以

$$\int_1^2 x^2\mathrm{d}x = \left.\frac{x^3}{3}\right|_1^2.$$

现在把 $x=2$ 和 $x=1$ 代入 $x^3/3$, 计算它们的差

$$\int_1^2 x^2\mathrm{d}x = \left.\frac{x^3}{3}\right|_1^2 = \left(\frac{2^3}{3}\right) - \left(\frac{1^3}{3}\right),$$

为 7/3. 还有另一个例子. 假设要计算

$$\int_{\pi/6}^{\pi/2} \cos(x)\mathrm{d}x.$$

我们需要知道 $\cos(x)$ 的反导数. 幸运的是, 我们知道一个反导数, 它是 $\sin(x)$. 毕竟, $\cos(x)$ 是 $\sin(x)$ 关于 x 的导数. 所以有

$$\int_{\pi/6}^{\pi/2} \cos(x)\mathrm{d}x = \sin(x)\Big|_{\pi/6}^{\pi/2} = \sin\left(\frac{\pi}{2}\right) - \sin\left(\frac{\pi}{6}\right) = 1 - \frac{1}{2} = \frac{1}{2}.$$

在之后的 17.6 节中, 我们会看到更多的例子.

17.4 不定积分

到目前为止, 我们使用两种不同的方法计算定积分: 黎曼和的极限 (太烦琐了) 和反导数 (不算太糟糕). 很显然, 我们不得不很熟练地寻找一个函数的反导数 —— 事实上, 在以后的章节中, 这种技巧非常必要. 所以, 可能需要一种简单的表示反导数的方式. 我们从微积分的第一基本定理中得到了灵感, 可以用

$$\int f(x)\mathrm{d}x$$

表示 "函数 f 的反导数的集合". 请记住任何可积函数都有无限多个反导数, 它们唯一的不同是常数部分. 这就是我说的 "集合" 的意思. 例如,

$$\int x^2\mathrm{d}x = \frac{x^3}{3} + C$$

对于任何常数 C 成立. 这个等式说明 x^2(关于 x) 的反导数是 $x^3/3+C$, 其中 C 是任意的常数. 如果忽略 C, 那么这个结果就是错误的, 因为这样只会得到一个反导数, 而实际上我们需要所有的反导数.

如果你知道一个函数的导数, 那么就会很快求出这个导数的反导数. 具体情况是:

$$\text{如果 } \frac{\mathrm{d}}{\mathrm{d}x}F(x) = f(x), \quad \text{那么} \quad \int f(x)\mathrm{d}x = F(x) + C.$$

上述例子适合这种情况：

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^3}{3}\right) = x^2, \quad \text{因此} \quad \int x^2\mathrm{d}x = \frac{x^3}{3} + C.$$

同样地, 有

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sin(x)) = \cos(x), \quad \text{因此} \quad \int \cos(x)\mathrm{d}x = \sin(x) + C.$$

到目前为止, 另一个例子为 (以后我们会有更多例子)：

$$\frac{\mathrm{d}}{\mathrm{d}x}(\tan^{-1}(x)) = \frac{1}{1+x^2}, \quad \text{因此} \quad \int \frac{1}{1+x^2}\mathrm{d}x = \tan^{-1}(x) + C.$$

再一次提醒, 常数 C 为任意常数. 本质上是任何可导函数只有一个导数, 而任何的可积函数都有无穷多个反导数.

上述的所有积分都是*不定积分*. 通过它们有无积分上下限, 可以区分定积分和不定积分. 不定积分没有积分上下限, 而定积分有. 这看起来可能是个很小的差别, 但实际上这两个积分有很大不同.

- 定积分, 如 $\int_a^b f(x)\mathrm{d}x$, 是一个数. 它表示由曲线 $y = f(x)$、x 轴以及垂线 $x = a$ 和 $x = b$ 所围成的面积.
- 不定积分, 如 $\int f(x)\mathrm{d}x$, 是一个函数的集合. 这个集合由函数 f 的所有反导数 (关于 x) 组成. 这些函数仅有的不同是它们的常数部分.

例如

$$\int_1^2 x^2\mathrm{d}x = \frac{7}{3}, \quad \text{而} \quad \int x^2\mathrm{d}x = \frac{x^3}{3} + C.$$

如果不是微积分的第二基本定理, 那么对于这两个表达式使用同样的符号 $\int$ 将是错误的. 幸运的是, 不定积分 (或者反导数) 正是你计算定积分所需要知道的东西, 所以我们在两个表达式中用了同样的符号.

这里有不定积分的两个性质, 它们来源于导数的相关性质：如果 f 和 g 是可积的, c 是一个常数, 这时

$$\int (f(x) + g(x))\mathrm{d}x = \int f(x)\mathrm{d}x + \int g(x)\mathrm{d}x$$

$$\int cf(x)\mathrm{d}x = c\int f(x)\mathrm{d}x.$$

也就是说, 和的积分是积分的和, 并且作为乘数的常数可以移到积分符号之外. 例如:

$$\int (5x^2 + 9\cos(x))\mathrm{d}x = 5\int x^2\mathrm{d}x + 9\int \cos(x)\mathrm{d}x = \frac{5x^3}{3} + 9\sin(x) + C.$$

注意我们仅仅需要一个常数 —— 尽管 $5x^3/3$ 和 $\sin(x)$ 都有它们自己的常数, 但你可以把这两个常数合并到一起. 顺便说一下, 适合于和的性质也适合于差:

$$\int (5x^2 - 9\cos(x))\mathrm{d}x = 5\int x^2\mathrm{d}x - 9\int \cos(x)\mathrm{d}x = \frac{5x^3}{3} - 9\sin(x) + C.$$

再一次提醒, 只需要一个常数.

在看其他例子之前, 我想对微积分的这两个基本定理再做些解释. 微积分的第一基本定理表明

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\mathrm{d}t = f(x).$$

从某种意义上讲, 积分的导数就是原始的函数. 你需要注意 "积分" 的意义, 请记住变量是积分上限而不是虚拟变量. 另外, 微积分的第二基本定理说明

$$\int_a^b f(x)\mathrm{d}x = F(x)\Big|_a^b,$$

其中 F 是 f 的反导数. 这就是说 $f(x) = \dfrac{\mathrm{d}}{\mathrm{d}x}F(x)$. 所以上述表达式可以重写为

$$\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}F(x)\mathrm{d}x = F(x)\Big|_a^b,$$

将其解释为一个函数的导数的积分就是这个函数本身. 再一次提醒, 它不是实际的原始函数, 它应该是原始函数在 a 和 b 两点数值的差. 这是很显然的, 微分和导数是相反的运算.

现在让我们看看怎样运用微积分的基本定理去解决问题.

17.5 怎样解决问题: 微积分的第一基本定理

思考一下怎样计算导数

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_3^x \sin(t^2)\mathrm{d}t.$$

你可以试着计算 $\int \sin(t^2)\mathrm{d}t$ 这个不定积分, 再把 x 和 3 代入求差, 这会得出

$$\int_3^x \sin(t^2)\mathrm{d}t.$$

最后再求这个结果的导数. 为什么不考虑积分和导数可以互相抵消的性质呢? 毕竟, 如果想要计算 $(\sqrt{54\,756})^2$, 没有必要浪费时间先计算 $\sqrt{54\,756}$ 的结果再去平方,

可以直接写下答案 54 756. 同样, 使用微积分的第一基本定理可以得到

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_3^x \sin(t^2)\mathrm{d}t = \sin(x^2).$$

所有需要做的是把被积函数 $\sin(t^2)$ 中的 t 改为 x. 数值 3 对我们的计算结果没有影响 (参见 17.1 节的讨论). 顺便说一下, 在计算结果后面放上 "$+C$" 是个严重的错误：你正在求导, 而不是反导!

当然, 你需要灵活掌握这个知识点 —— 用任何字母来表示变量. 例如,

$$\frac{\mathrm{d}}{\mathrm{d}z}\int_{-\mathrm{e}}^z 2^{\cos(w^2\ln(w+5))}\mathrm{d}w$$

是什么意思? 我们可以用 z 替代被积函数中的 w, 这样有

$$\frac{\mathrm{d}}{\mathrm{d}z}\int_{-\mathrm{e}}^z 2^{\cos(w^2\ln(w+5))}\mathrm{d}w = 2^{\cos(z^2\ln(z+5))}.$$

注意 $-\mathrm{e}$ 是一个常数, 但再一次提醒的是, 它可以用任何其他常数替代, 而答案会是一样的. (顺便说一下, 该积分只有当 $z > -5$ 时才有意义.)

这同上一节的讲解是一样的, 该函数的变量 (也就是对谁求导) 就是积分上限. 所有需要做的是用实际的变量去替代虚拟变量. 还有其他四种情况, 我们分别看看.

17.5.1 变形 1：变量是积分下限

考虑积分

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_x^7 t^3\cos(t\ln(t))\mathrm{d}t.$$

问题是变量是积分下限而不是积分上限. 没问题, 只要把 x 和 7 互换, 再在新的积分前面加个负号 (参见 16.3 节计算这个积分). 可以得到

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_x^7 t^3\cos(t\ln(t))\mathrm{d}t = \frac{\mathrm{d}}{\mathrm{d}x}\left(-\int_7^x t^3\cos(t\ln(t))\mathrm{d}t\right).$$

现在把负号移出积分符号, 然后使用微积分的第一基本定理解决这个问题. 如果 $x > 0$, 该结果为

$$-x^3\cos(x\ln(x)),$$

实际上, 我们所要做的就是提取被积函数, 用变量 x 替代虚拟变量 t, 再在前面加上负号. 在前面加上负号, 再互换积分上下限使用微积分的第一基本定理, 这很重要, 就像在前面例子中见过的那样.

17.5.2 变形 2：积分上限是一个函数

这有另一个例子：

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^{x^2}\tan^{-1}(t^7+3t)\mathrm{d}t.$$

因为积分上限是 x^2 而不是 x, 所以不能直接使用微积分的第一基本定理, 需要使用链式求导法则. 我们可以设置这个积分为 y, 然后再求导:

$$y=\int_0^{x^2}\tan^{-1}(t^7+3t)\mathrm{d}t.$$

我们想要计算 $\mathrm{d}y/\mathrm{d}x$. 因为 y 是一个关于 x^2 的函数, 而不直接关于 x, 我们可以设置 $u=x^2$. 这就是说

$$y=\int_0^{u}\tan^{-1}(t^7+3t)\mathrm{d}t.$$

链式法则告诉我们,

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x};$$

而第一基本定理告诉我们,

$$\frac{\mathrm{d}y}{\mathrm{d}u}=\frac{\mathrm{d}}{\mathrm{d}u}\int_0^{u}\tan^{-1}(t^7+3t)\mathrm{d}t=\tan^{-1}(u^7+3u).$$

又因为 $u=x^2$, 所以有 $\mathrm{d}u/\mathrm{d}x=2x$. 这样,

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}=(\tan^{-1}(u^7+3u))(2x).$$

现在需要做的是用 x^2 替代 u, 可得

$$\frac{\mathrm{d}y}{\mathrm{d}x}=2x\tan^{-1}((x^2)^7+3(x^2))=2x\tan^{-1}(x^{14}+3x^2).$$

即,

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^{x^2}\tan^{-1}(t^7+3t)\mathrm{d}t=2x\tan^{-1}(x^{14}+3x^2).$$

当你把它分解时, 这并不太糟.



让我们看看这种问题的另一个例子:

$$\frac{\mathrm{d}}{\mathrm{d}q}\int_4^{\sin(q)}\tan(\cos(a))\mathrm{d}a$$

是怎么回事儿呢? 设积分

$$y=\int_4^{\sin(q)}\tan(\cos(a))\mathrm{d}a,$$

并时刻提醒自己正在计算 $\mathrm{d}y/\mathrm{d}q$. 现在设 $u=\sin(q)$, 所以

$$y=\int_4^{u}\tan(\cos(a))\mathrm{d}a.$$

通过链式求导法则, 有

$$\frac{\mathrm{d}y}{\mathrm{d}q}=\frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}q}.$$

根据微积分的第一基本定理：

$$\frac{\mathrm{d}y}{\mathrm{d}u}=\frac{\mathrm{d}}{\mathrm{d}u}\int_4^u \tan(\cos(a))\mathrm{d}a=\tan(\cos(u)).$$

因为 $u=\sin(q)$, 所以有 $\mathrm{d}u/\mathrm{d}q=\cos(q)$, 上述积分应用链式法则后为

$$\frac{\mathrm{d}u}{\mathrm{d}q}=\frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}q}=\tan(\cos(u))\cos(q).$$

最后用 $\sin(q)$ 替代 u, 这样有

$$\frac{\mathrm{d}}{\mathrm{d}q}\int_4^{\sin(q)} \tan(\cos(a))\mathrm{d}a=\tan(\cos(\sin(q)))\cos(q).$$

你可能在同一问题中遇到过上述两种情况. 例如, 计算

$$\frac{\mathrm{d}}{\mathrm{d}q}\int_{\sin(q)}^{4} \tan(\cos(a))\mathrm{d}a$$

时, 可以首先交换积分上下限, 在前面加上一个减号, 这样有

$$\frac{\mathrm{d}}{\mathrm{d}q}\int_{\sin(q)}^{4} \tan(\cos(a))\mathrm{d}a=-\frac{\mathrm{d}}{\mathrm{d}q}\int_4^{\sin(q)} \tan(\cos(a))\mathrm{d}a.$$

现在积分上限像我们上道例题那样. 最后的极限结果会是一样的, 只是多了一个负号：

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\int_{\sin(q)}^{4} \tan(\cos(a))\mathrm{d}a&=-\frac{\mathrm{d}}{\mathrm{d}q}\int_4^{\sin(q)} \tan(\cos(a))\mathrm{d}a\\&=-\tan(\cos(\sin(q))\cos(q).\end{aligned}$$

17.5.3 变形 3：积分上下限都为函数

这是另一个更为复杂的例子：

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{x^5}^{x^6} \ln(t^2-\sin(t)+7)\mathrm{d}t,$$

其积分上下限都是关于 x 的函数. 解决这个问题的方法是用一个常数把这个积分分成两个部分. 在哪里分开这个积分并不重要, 只要所使用的常数在该被积函数的定义域内. 所以, 选一个你最喜欢的数 —— 我们选 0, 这样把该积分分为：

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}x}\int_{x_5}^{x^6}\ln(t^2-\sin(t)+7)\mathrm{d}t\\&=\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_{x^5}^{0}\ln(t^2-\sin(t)+7)\mathrm{d}t+\int_{0}^{x^6}\ln(t^2-\sin(t)+7)\mathrm{d}t\right).\end{aligned}$$

这样, 便把这个问题分解成了两个简单的导数. 前面这个积分就是前面两种情况的混合. 通过交换积分上下限, 并在前面加上负号, 我们有

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{x^5}^{0}\ln(t^2-\sin(t)+7)\mathrm{d}t=-\frac{\mathrm{d}}{\mathrm{d}x}\int_{0}^{x^5}\ln(t^2-\sin(t)+7)\mathrm{d}t.$$



现在通过设 $u=x^5$ 使用链式求导法则, 然后用上一节的方法. 计算后会得出这个导数为

$$-5x^4\ln((x^5)^2-\sin(x^5)+7)=-5x^4\ln(x^{10}-\sin(x^5)+7).$$

这个导数的另一部分是

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{0}^{x^6}\ln(t^2-\sin(t)+7)\mathrm{d}t,$$

这次我们不用交换积分上下限了 —— 仅仅设 $v=x^6$ 然后再次应用链式求导法则. 你会发现上述的导数等于

$$6x^5\ln((x^6)^2-\sin(x^6)+7)=6x^5\ln(x^{12}-\sin(x^6)+7).$$

再将两者放在一起, 有

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}x}\int_{x^5}^{x^6}\ln(t^2-\sin(t)+7)\mathrm{d}t\\&=-5x^4\ln(x^{10}-\sin(x^5)+7)+6x^5\ln(x^{12}-\sin(x^6)+7).\end{aligned}$$

17.5.4　变形 4：导数伪装成极限



这是一个看起来很不同的例子：

$$\lim_{h\to 0}\frac{1}{h}\int_{x}^{x+h}\log_3(\cos^6(t)+2)\mathrm{d}t.$$

这不是一个导数, 它是一个极限. 实际上, 它是伪装的导数 (参见 6.5 节关于这种极限的讨论). 技巧是对于某个常数 a 设

$$F(x)=\int_{a}^{x}\log_3(\cos^6(t)+2)\mathrm{d}t.$$

也可以用一个指定的常数, 或则干脆就使用 a. 这都无关紧要, 因为在任何情况下, 我们都有

$$F(x+h)-F(x)=\int_x^{x+h}\log_3(\cos^6(t)+2)\mathrm{d}t.$$

如果你不信, 可以自己校验一下; 也可参见 17.2 节. 在任何情况下, 关于函数 F, 我们都有

$$\lim_{h\to0}\frac{1}{h}\int_x^{x+h}\log_3(\cos^6(t)+2)\mathrm{d}t=\lim_{h\to0}\frac{F(x+h)-F(x)}{h}=F'(x).$$

所以, 实际上对于任何常数 a, 我们有

$$\lim_{h\to0}\frac{1}{h}\int_x^{x+h}\log_3(\cos^6(t)+2)\mathrm{d}t=\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x\log_3(\cos^6(t)+2)\mathrm{d}t.$$

看, 我告诉过你, 这个极限是个伪装的导数! 为了解决这个问题, 我们可以使用微积分的第一基本定理, 通过计算可知该极限为 $\log_3(\cos^6(x)+2)$.

17.6　怎样解决问题: 微积分的第二基本定理

使用微积分的第二基本定理计算定积分 (这是计算定积分的方法, 相信我) 首先要找到不定积分, 然后分别把积分上下限代入, 最后再求差. 所以让我们花一些时间讨论怎样找到不定积分 (也就是反导数), 然后再看一些计算定积分的例子. 这只是积分学的开始, 在后两章中, 我们将会看到更多计算不定积分的方法.

17.6.1　计算不定积分

像我们在 17.4 节中看到的, 只要知道一个函数的导函数, 那么就一定会知道这个导函数的反导数. 我们已经给出了一些例子, 这里还有另一个: 因为

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^4)=4x^3,$$

所以立刻可知

$$\int 4x^3\mathrm{d}x=x^4+C.$$

因为常数可以被移到积分符号的外边, 所以改写为

$$4\int x^3\mathrm{d}x=x^4+C.$$

现在两边分别除以 4:

$$\int x^3\mathrm{d}x=\frac{x^4}{4}+\frac{C}{4}.$$

这很好, 但 $C/4$ 看上去有些傻. 任意常数除以 4 得到的还是任意常数. 所以可以用任意常数去替代 $C/4$, 我们还使用 C, 这样有

$$\int x^3\mathrm{d}x=\frac{x^4}{4}+C.$$

让我们对 x 的幂重复这个计算. 注意

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^{a+1}) = (a+1)x^a;$$

这就是说

$$\int (a+1)x^a \mathrm{d}x = x^{a+1} + C.$$

如果 $a \neq -1$, 这时 $a+1 \neq 0$; 所以可以等式两边同时除以 $(a+1)$, 把它写为

$$\boxed{\int x^a \mathrm{d}x = \frac{x^{a+1}}{a+1} + C.}$$

(再一次提醒, 我们用 C 替代了 $C/(a+1)$; 这是可以的, 因为 C 仅仅是任意常数.) 那么当 $a=-1$ 时, 情况又是怎样呢? 上述的方法并不适用于

$$\int \frac{1}{x} \mathrm{d}x.$$

另外, 从 9.3 节可知

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln(x)) = \frac{1}{x}, \quad 因此 \quad \int \frac{1}{x} \mathrm{d}x = \ln(x) + C.$$

这很好, 但实际上我们可以做得更好. 你看, 除了在 $x = 0$ 点外, $1/x$ 在任意一点都有意义, 而 $\ln(x)$ 仅仅当 $x > 0$ 时才有意义. 我们可以这样改写来弥补这个不足:

$$\int \frac{1}{x} \mathrm{d}x = \ln|x| + C.$$

让我们检查一下刚才的计算是否正确. 我们需要证明

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln|x| = \frac{1}{x}$$

对于所有的 $x \neq 0$ 都成立. 当 $x > 0$ 时, 左边就是 $\ln(x)$, 符合要求; 当 $x < 0$ 时, $|x|$ 实际上等于 $-x$, 所以这时左边为

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln(-x).$$

它看起来很奇怪, 但请记住当 $x < 0$ 时, $-x$ 为正. 在这种情况下, 通过链式求导法则, 上述的导数为

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln(-x) = -\frac{1}{-x} = \frac{1}{x}.$$

所以我们已经证明了:

$$\boxed{\int \frac{1}{x} \mathrm{d}x = \ln|x| + C.}$$

参见 17.7 节关于使用这个公式的技巧. 与此同时, 我们要用基本的求导公式总结相应的积分公式.

导数和积分公式

$$\frac{\mathrm{d}}{\mathrm{d}x}x^a = ax^{a-1} \qquad \int x^a\mathrm{d}x = \frac{x^{a+1}}{a+1} + C \quad (\text{如果 } a \neq -1)$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln(x) = \frac{1}{x} \qquad \int \frac{1}{x}\mathrm{d}x = \ln|x| + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^x = \mathrm{e}^x \qquad \int \mathrm{e}^x\mathrm{d}x = \mathrm{e}^x + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}b^x = b^x\ln(b) \qquad \int b^x\mathrm{d}x = \frac{b^x}{\ln(b)} + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \cos(x) \qquad \int \cos(x)\mathrm{d}x = \sin(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = -\sin(x) \qquad \int \sin(x)\mathrm{d}x = -\cos(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan(x) = \sec^2(x) \qquad \int \sec^2(x)\mathrm{d}x = \tan(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sec(x) = \sec(x)\tan(x) \qquad \int \sec(x)\tan(x)\mathrm{d}x = \sec(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cot(x) = -\csc^2(x) \qquad \int \csc^2(x)\mathrm{d}x = -\cot(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\csc(x) = -\csc(x)\cot(x) \qquad \int \csc(x)\cot(x)\mathrm{d}x = -\csc(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin^{-1}(x) = \frac{1}{\sqrt{1-x^2}} \qquad \int \frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \sin^{-1}(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan^{-1}(x) = \frac{1}{1+x^2} \qquad \int \frac{1}{1+x^2}\mathrm{d}x = \tan^{-1}(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sec^{-1}(x) = \frac{1}{|x|\sqrt{x^2-1}} \qquad \int \frac{1}{|x|\sqrt{x^2-1}}\mathrm{d}x = \sec^{-1}(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sinh(x) = \cosh(x) \qquad \int \cosh(x)\mathrm{d}x = \sinh(x) + C$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cosh(x) = \sinh(x) \qquad \int \sinh(x)\mathrm{d}x = \cosh(x) + C$$

如我们所知, 在上述微分公式中, 如果用 ax 替代 x, 那么把每一个相应的公式乘以 a 就可以了. 例如:

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan(7x) = 7\sec^2(7x).$$

但如果是积分呢? 现在这个规则是这样的: 如果你用 ax 替代 x, 这时需要把相应的公式除以 a. 例如:



$$\int \sec^2(7x)\mathrm{d}x = \frac{1}{7}\tan(7x) + C.$$

从这个例子被 7 除可以直接看出这个说法是正确的. 这有另一个例子:

$$\int \mathrm{e}^{-x/3}\mathrm{d}x.$$

你可以把 x 看作被 $-1/3$ 倍的 x 替代; 所以除以 $-1/3$ 可得

$$\int \mathrm{e}^{-x/3}\mathrm{d}x = \frac{1}{-1/3}\mathrm{e}^{-x/3} + C = -3\mathrm{e}^{-x/3} + C.$$

再多练习一个怎么样? 考虑

$$\int \frac{1}{1+2x^2}\mathrm{d}x.$$

这个积分可以改写为

$$\int \frac{1}{1+(\sqrt{2}x)^2}\mathrm{d}x,$$

现在可以把 x 看作被 $\sqrt{2}x$ 替代. 所以除以 $\sqrt{2}$ 可得

$$\int \frac{1}{1+(\sqrt{2}x)^2}\mathrm{d}x = \frac{1}{\sqrt{2}}\tan^{-1}(\sqrt{2}x) + C.$$

在后两章中, 我们将会看到更多更复杂的计算反导数的技巧, 但也要先记住这个简单的, 因为常数作倍数是积分中常见的现象.

17.6.2 计算定积分

微积分的第二基本定理告诉我们, 为计算

$$\int_a^b f(x)\mathrm{d}x,$$

仅仅需要先找到它的反导数, 然后把 $x = a$ 和 $x = b$ 分别代入, 最后求它们的差. 在 17.3 节中, 我们已经看到了一些例子, 现在再看 5 个例子. 首先, 考虑

$$\int_{-1}^{2} x^4\mathrm{d}x.$$

通过使用公式

$$\int x^a\mathrm{d}x = \frac{x^{a+1}}{a+1} + C,$$

我们知道 x^4 的反导数是 $x^5/5$. 没有必要考虑这个常数, 你可以选择任何反导数, 我们简单地选取 $C = 0$ 这个反导数. 所以有

$$\int_{-1}^{2} x^4\mathrm{d}x = \left.\frac{x^5}{5}\right|_{-1}^{2} = \left(\frac{2^5}{5}\right) - \left(\frac{(-1)^5}{5}\right) = \left(\frac{32}{5}\right) - \left(\frac{-1}{5}\right) = \frac{33}{5}.$$

使用括号很重要, 因为这样可以避免丢掉负号! 现在你可能会考虑, 如果我们使用不同的反导数情况会是怎样. 这个想法很好, 但常数最后是会被抵消的. 例如, 如果你选 $x^5/5 - 1001$ 作为它的反导数, 这样会得到

$$\int_{-1}^{2} x^4 \mathrm{d}x = \left(\frac{x^5}{5} - 1001\right)\Bigg|_{-1}^{2} = \left(\frac{2^5}{5} - 1001\right) - \left(\frac{(-1)^5}{5} - 1001\right)$$
$$= \left(\frac{2^5}{5}\right) - 1001 - \left(\frac{(-1)^5}{5}\right) + 1001.$$

注意 -1001 和 $+1001$ 这两项相互抵消了, 我们得到的正是之前的结果. 这个方法给我们的启迪是, 当计算定积分时可以忽略常数 C.

第二个例子是：

$$\int_{-\mathrm{e}^2}^{-1} \frac{4}{x} \mathrm{d}x.$$

常数 4 可以移到积分符号的外边, 所以我们需要使用公式

$$\int \frac{1}{x} \mathrm{d}x = \ln|x| + C.$$

从前述总结的公式表中可以看出, $4\ln|x|$ 是 $4/x$ 的反导数. 所以有

$$\int_{-\mathrm{e}^2}^{-1} \frac{4}{x} \mathrm{d}x = 4\ln|x|\Bigg|_{-\mathrm{e}^2}^{-1} = (4\ln|-1|) - (4\ln|-\mathrm{e}^2|) = 4\ln(1) - 4\ln(\mathrm{e}^2) = -8.$$

在这儿, 我们使用了 $\ln(1) = 0, \ln(\mathrm{e}^2) = 2\ln(\mathrm{e}) = 2$.

第三个例子是

$$\int_{0}^{\pi/3} \left(\sec^2(x) - 5\sin\left(\frac{x}{2}\right)\right) \mathrm{d}x.$$

你应该马上就能看出来, 应该把这个积分分成两部分：$\sec^2(x)$ 和 $\sin(x/2)$, 不考虑第二个积分外面的常数. 根据公式表可得, $\sec^2(x)$ 的反导数为 $\tan(x)$; 对于 $\sin(x/2)$, 它的反导数是 $-\cos(x/2)$ 除以 $1/2$, 因为 x 可以被它的常数倍 $x/2$ 所替代. 这样结果为 $-2\cos(x/2)$(因为除以 $1/2$ 和乘以 2 是等价的). 综合在一起, 我们有

$$\int_{0}^{\pi/3} \left(\sec^2(x) - 5\sin\left(\frac{x}{2}\right)\right) \mathrm{d}x = \left(\tan(x) - 5 \times \left(-2\cos\left(\frac{x}{2}\right)\right)\right)\Bigg|_{0}^{\pi/3}.$$

通过化简和替代有

$$\left(\tan(\pi/3) + 10\cos\left(\frac{\pi/3}{2}\right)\right) - \left(\tan(0) + 10\cos\left(\frac{0}{2}\right)\right);$$

你可以发现最终结果为 $6\sqrt{3} - 10$.

这是第四个例子：

$$\int_{4}^{9} \frac{1}{x\sqrt{x}} \mathrm{d}x.$$

解这道题的技巧是把被积函数写为 $x^{-3/2}$ 的形式. 确信你理解这种写法! 现在我们可以使用公式表里的 $\int x^a \mathrm{d}x$ 解决这个问题.

$$\int_4^9 \frac{1}{x\sqrt{x}}\mathrm{d}x = \int_4^9 x^{-3/2}\mathrm{d}x = \frac{1}{-1/2}x^{-1/2}\Big|_4^9 = (-2\times(9)^{-1/2}) - (-2\times(4)^{-1/2})$$
$$= -\frac{2}{3}+\frac{2}{2}=\frac{1}{3}.$$

这节的最后一个例子为

$$\int_0^{1/6}\frac{\mathrm{d}x}{\sqrt{1-9x^2}}.$$

不要因为把 $\mathrm{d}x$ 写在分子的位置就看不懂了, 其实就是换了一个写法, 它与下式等价:

$$\int_0^{1/6}\frac{1}{\sqrt{1-9x^2}}\mathrm{d}x.$$

我们用 $(3x)^2$ 替代 $9x^2$, 这样有

$$\int_0^{1/6}\frac{\mathrm{d}x}{\sqrt{1-9x^2}} = \int_0^{1/6}\frac{1}{\sqrt{1-(3x)^2}}\mathrm{d}x = \frac{1}{3}\sin^{-1}(3x)\Big|_0^{1/6}.$$

从上面的公式表可知

$$\int\frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \sin^{-1}(x)+C.$$

但我们需要除以 3, 因为 x 被 $3x$ 替代. 现在计算这个定积分的值:

$$\left(\frac{1}{3}\sin^{-1}\left(3\times\frac{1}{6}\right)\right) - \left(\frac{1}{3}\sin^{-1}(3\times 0)\right) = \left(\frac{1}{3}\times\frac{\pi}{6}\right) - (0) = \frac{\pi}{18}.$$

这里我们利用了 $\sin^{-1}\left(\dfrac{1}{2}\right)=\pi/6$.

17.6.3 面积和绝对值

在 16.1 节中, 我们见过

$$\int_{-\pi}^{\pi}\sin(x)\mathrm{d}x = 0,$$

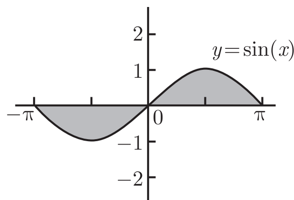


图 17-5

因为坐标轴上下的面积可以互相抵消. 图 17-5 是这个积分的图像.

我们可以用反导数的方法来计算这个定积分:

$$\int_{-\pi}^{\pi}\sin(x)\mathrm{d}x = -\cos(x)\Big|_{-\pi}^{\pi} = (-\cos(\pi)) - (-\cos(-\pi))$$
$$= -(-1)+(-1)=0.$$

如果不考虑面积的正负, 也就是只计算实际面积, 那么刚才的例题又是怎样呢? 在 16.4.1 节中, 我们看到了解决这个问题的例子: 这个以平方为单位的面积等于

$$\int_{-\pi}^{\pi}|\sin(x)|\mathrm{d}x.$$

我们的方法是把这个原始积分在它与 x 轴的交点处分成两部分, 这时再取每一部分的绝对值：

$$\int_{-\pi}^{\pi} |\sin(x)|\mathrm{d}x = \left|\int_{-\pi}^{0} \sin(x)\mathrm{d}x\right| + \left|\int_{0}^{\pi} \sin(x)\mathrm{d}x\right|.$$



我们可以使用它的反导数 $-\cos x$ 计算这两个积分的结果, 它们分别为 -2 和 2, 我把这个计算工作留给你. 如果简单地把这两个数加到一起, 就得到了有向面积为 0 平方单位; 但如果首先取绝对值, 那么就可得到这个面积的实际值, 它是 $|-2|+|2|=4$ 平方单位.



现在来看看求两条曲线间的面积的例子. 我们在 16.4.2 节中已经说明该怎样计算, 但现在可以使用微积分的第二基本定理这个强大的工具来帮助我们解决问题. 我们可以求图 17-6 所示的不规则图形的面积.

我们计算由 $y=x$、$y=1/x$ 和直线 $x=2$ 所围成的面积. 需要找到 $y=x$ 和 $y=1/x$ 的交点：设 $x=1/x$ 可以得 $x^2=1$; 也就是说, $x=1$ 或 $x=-1$. 在这个图像中, 交点的横坐标为正, 所以我们选择 $x=1$. 因为 $y=x$ 在 $y=1/x$ 的上边, 我们用上边的函数减下边的函数并求积分可得

$$\text{阴影部分面积} = \int_{1}^{2}\left(x-\frac{1}{x}\right)\mathrm{d}x.$$

可以容易地使用 $\int x^a \mathrm{d}x = x^{a+1}/(a+1)+C$ 求 x 的反导数, 当 $a=1$ 时, 该反导数为 $x^2/2$; 并且我们也知道, $1/x$ 的反导数为 $\ln|x|$. 所以上述积分等于

$$\left(\frac{x^2}{2}-\ln|x|\right)\Bigg|_1^2 = \left(\frac{2^2}{2}-\ln|2|\right)-\left(\frac{1^2}{2}-\ln|1|\right) = 2-\ln(2)-\frac{1}{2}+\ln(1).$$

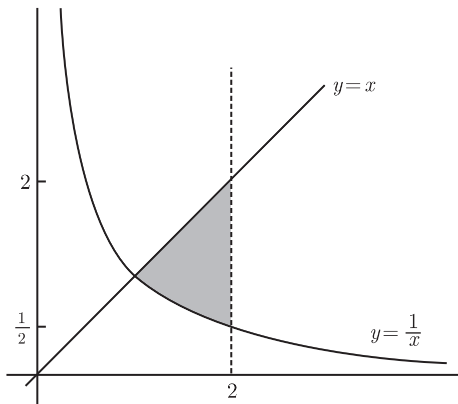


图　17-6

化简后为 $3/2-\ln(2)$, 即我们要计算的这块面积为 $3/2-\ln(2)$ 平方单位. 现在让我们看看图 17-7, 如果计算这个面积该怎样做?

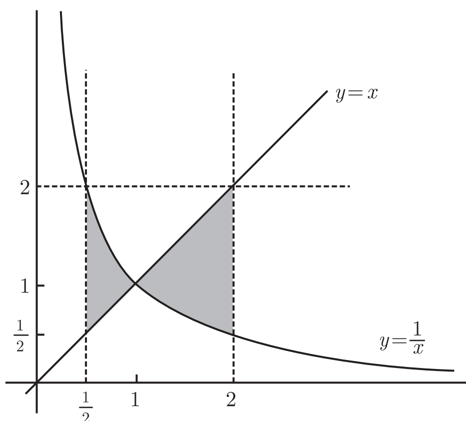


图 17-7

让我们试着把这个面积写为

$$\text{新的阴影部分面积} \stackrel{?}{=} \int_{1/2}^{2} \left(x - \frac{1}{x}\right) \mathrm{d}x,$$

但实际上这是不对的. 你看, 在区间 1/2 和 1 之间曲线 $y = x$ 不在函数 $y = 1/x$ 的上边. 在 16.4.2 节中我们讨论过这个问题, 实际上需要取被积函数的绝对值:

$$\text{新的阴影部分面积} = \int_{1/2}^{2} \left|x - \frac{1}{x}\right| \mathrm{d}x.$$

因为唯一的交点是在 $x = 1$ 处, 所以我们从该点分割这个积分, 然后分别取绝对值再求积分

$$\int_{1/2}^{2} \left|x - \frac{1}{x}\right| \mathrm{d}x = \left|\int_{1/2}^{1} \left(x - \frac{1}{x}\right) \mathrm{d}x\right| + \left|\int_{1}^{2} \left(x - \frac{1}{x}\right) \mathrm{d}x\right|.$$

我们已经求过第二个积分, 它的结果为 $3/2 - \ln(2)$, 因为 $\ln(2) < \ln(\mathrm{e}) = 1$, 所以这个值是正的. 对于第一个积分, 我们有

$$\begin{aligned}
\int_{1/2}^{1} \left(x - \frac{1}{x}\right) \mathrm{d}x &= \left(\frac{x^2}{2} - \ln|x|\right) \Big|_{1/2}^{1} \\
&= \left(\frac{1^2}{2} - \ln|1|\right) - \left(\frac{(1/2)^2}{2} - \ln\left|\frac{1}{2}\right|\right) \\
&= \frac{1}{2} - \ln(1) - \frac{1}{8} + \ln\left(\frac{1}{2}\right) = \frac{3}{8} - \ln(2).
\end{aligned}$$

在这里可以使用 9.1.4 节中的对数法则, 把 $\ln(1/2)$ 改写为 $\ln(1/2) = \ln(1) - \ln(2)$ 或 $\ln(1/2) = \ln(2^{-1})$, 这样就可以说 $\ln(1/2) = -\ln(2)$. 请注意 $3/8 - \ln(2)$ 的值为负. 当在 $[1/2, 1]$ 这个区间时, x 是比 $1/x$ 小的, 所以 $x - 1/x$ 的积分为负. 我们取 $3/8 - \ln(2)$ 的绝对值, 即 $\ln(2) - 3/8$. 所以有

$$\left|\int_{1/2}^{1}\left(x-\frac{1}{x}\right)\mathrm{d}x\right|+\left|\int_{1}^{2}\left(x-\frac{1}{x}\right)\mathrm{d}x\right|=\left|\frac{3}{8}-\ln(2)\right|+\left|\frac{3}{2}-\ln(2)\right|$$
$$=\left(\ln(2)-\frac{3}{8}\right)+\left(\frac{3}{2}-\ln(2)\right)=\frac{9}{8}.$$

我们要计算的阴影部分面积是 9/8 平方单位. 实际上计算这个面积可以不用微积分的方法. 请看图, 我们可以发现 $y=x$ 和 $y=1/x$ 关于直线 $y=x$ 对称, 所以如果把这个楔形物移动到直线 $y=x$ 的上边, 这时它组成了一个三角形, 如图 17-8 所示.

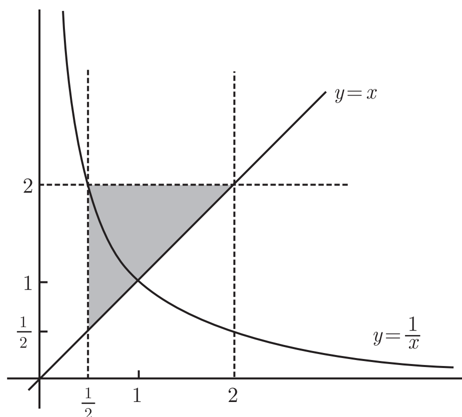


图　17-8

这个三角形的底和高都是 3/2 单位, 所以它的面积为 9/8 平方单位, 同我们刚才计算的结果是一致的!

17.7　技 术 要 点



在 17.6.1 节, 我们知道

$$\int\frac{1}{x}\mathrm{d}x=\ln|x|+C.$$

尽管每个人都这样写这个公式, 但从技术上说这并不正确. 你知道, 我们想要求所有的 $1/x$ 的反导数. 虽然对于每个不同的常数 C, $\ln|x|+C$ 都是它的一个反导数, 但实际上还有更多. 要知道原因, 让我们看看函数 $y=\ln|x|$ 的图像, 如图 17-9 所示.

这个图像有两部分, 我们可以任意上下移动其中的一部分, 却不影响它的导数的结果. 例如, 如果把左边的图像向上移动一个单位, 把右边的图像向下移动 1/2 个单位, 图像将会如图 17-10 所示.

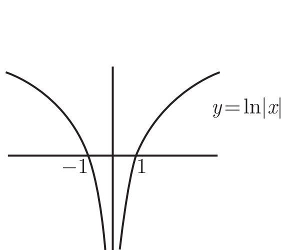


图 17-9

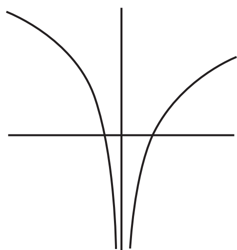


图 17-10

该函数不是 $\ln|x|+C$ 这种形式, 但是它的导数仍然是 $1/x$. 所以我们真的需要两个常数项, 这两项是不同的, 每一项对应这两个函数中的一个：

$$\int \frac{1}{x}\mathrm{d}x = \begin{cases} \ln|x|+C_1 & \text{如果 } x<0, \\ \ln|x|+C_2 & \text{如果 } x>0. \end{cases}$$

我们通常只写一个而不写两个常数的原因是, 在一次计算中只用到了一个常数. 考虑以下三个积分：

$$\int_1^{\mathrm{e}} \frac{1}{x}\mathrm{d}x, \quad \int_{-\mathrm{e}}^{-1} \frac{1}{x}\mathrm{d}x, \quad \int_{-1}^{\mathrm{e}} \frac{1}{x}\mathrm{d}x.$$

在第一个积分中, 我们只用到了 $y=1/x$ 图像的右侧分支. 同样, 我们对第二个积分只用了图像的左侧分支. 通过计算可得, 这两个积分的结果分别为 1 和 -1. 至于第三个积分, 我们需要使用这个图像的两个分支了, 但出问题了：在区间 $[-1, \mathrm{e}]$ 有一条垂直渐近线 $x=0$, 我们不知道怎样去做. 事实上, 在第 20 章的反常积分中, 我们将学习如何处理这种类型的积分. 而本例中, 由于这条垂直渐近线使得第三个积分看起来没有意义. 所以对于

$$\int_a^b \frac{1}{x}\mathrm{d}x,$$

这种类型的定积分有意义的情况是当 a 和 b 同时为正或同时为负. 在任何一种情况下, 我们只需要使用其中的一个分支, 没有必要考虑两个常数的问题了.

17.8 微积分第一基本定理的证明

在 17.2 节中, 我们给出了微积分第一基本定理的大致证明. 现在, 我们要使这个定理更加严谨. 回顾这个式子：

$$F(x) = \int_a^x f(t)\mathrm{d}t,$$

我们要计算 $F'(x)$. 我们已经看到

$$F(x+h)-F(x)=\int_x^{x+h} f(t)\mathrm{d}t.$$

假设 $h>0$. 根据积分学的中值定理有 (参见 16.6.1 节), 在区间 $[x, x+h]$ 上有个常数 c 使得

$$\int_x^{x+h} f(t)\mathrm{d}t=((x+h)-x)f(c)$$

成立. 这样, 我们有

$$F(x+h)-F(x)=\int_x^{x+h} f(t)\mathrm{d}t=hf(c)$$

对于某个在区间 $[x, x+h]$ 内的常数 c 成立. 实际上对于 $h<0$, 这个等式也是成立的, 这时的区间就变为 $[x+h, x]$, 因为在这种情况下 $x+h<x$. 而等式的两端同时除以 h 有

$$\frac{F(x+h)-F(x)}{h}=f(c).$$

关键点是：当 x 是一个固定数时 (暂时), 数 c 的变化是由 h 决定的, 而且它在 x 和 $x+h$ 之间. 可能我们需要重写这个方程为：

$$\frac{F(x+h)-F(x)}{h}=f(c_h),$$

这样写就强调了 c 是由 h 决定的. 当 $h\to 0$ 时情况又怎样呢? 这个值 c_h 被夹在了 x 和 $x+h$ 之间, 所以当 $h\to 0$ 时, 根据三明治定理 (参见 3.6 节) 我们有 $c_h\to x$. 另一方面, 因为函数 f 是连续的, 当 $h\to 0$ 时, 一定有 $f(c_h)\to f(x)$. 也就是说,

$$\lim_{h\to 0}\frac{F(x+h)-F(x)}{h}=\lim_{h\to 0} f(c_h)=f(x).$$

这足以说明 $F'(x)=f(x)$, 这就完成了这个定理的证明. 至于微积分的第二基本定理, 实际上已经在 17.3 节中证明过了, 所以我们可以进入下一章的学习了!

第 18 章　积分的方法 I

让我们开始发展一套求反导数的技巧. 在这一章中, 我们将要学习以下方法:

- 换元法 (也可以叫变量替换);
- 分部积分法;
- 使用部分分式对有理函数求积分.

在下一章中, 我们将会看到关于三角函数的更多技巧.

18.1　换　元　法

使用链式求导法则, 我们可以很容易地求出 e^{x^2} 关于 x 的导数, 请看

$$\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{x^2}) = 2x\mathrm{e}^{x^2}.$$

因子 $2x$ 是出现在指数位置的 x^2 的导数. 现在, 像我们在 17.4 节中见过的, 可以得到

$$\int 2x\mathrm{e}^{x^2}\mathrm{d}x = \mathrm{e}^{x^2} + C$$

对于任意常数 C 成立. 所以我们求得了 $2x\mathrm{e}^{x^2}$ 关于 x 的积分. 那么 e^{x^2} 的积分为多少呢？你可能认为求解积分

$$\int \mathrm{e}^{x^2}\mathrm{d}x$$

是很容易的. 这个积分好像并不难求, 但这是不可能的! 虽然不是完全不可能, 但事实上 e^{x^2} 的反导数表达式是很复杂的. (要计算这个积分你需要求助无穷级数、定积分或一些其他方法.) 你会不会认为, $\mathrm{e}^{x^2}/2x$ 为这个函数的反导数呢？不是. 你可以使用除法规则对这个函数求导 (关于 x), 求导后的结果和 e^{x^2} 有天壤之别.

能让我们解出 $\int 2x\mathrm{e}^{x^2}\mathrm{d}x$, 是由于 $2x$ 这个因子的存在, 该因子恰恰就是链式求导法则之后的 x^2 的导数. 现在考虑从下面的不定积分开始:



$$\int x^2\cos(x^3)\mathrm{d}x.$$

要求这个带着 x^3 的余弦函数的反导数, 我们有一线希望：里面的 x^3 的导数是 $3x^2$. 这几乎和被积函数里的一个因子 x^2 相匹配 —— 这里仅仅是常数 3 使得问题看起来有些难了. 但是常数可以移到积分符号的外边去, 所以这并不是一个问题.

让我们从设 $t = x^3$ 开始, 所以 $\cos(x^3)$ 变为 $\cos(t)$. 我们的目的是：要用 t 替

代表达式中的每一个 x. 你可能会说上述积分是变量 x 统治的领地, 但我们要把它变成 t 的领地. 我们已经把 $\cos(x^3)$ 替换掉, 但还需要考虑替换 x^2 和 $\mathrm{d}x$.

事实上, $\mathrm{d}x$ 是很重要的. 你不能随便地把它改为 $\mathrm{d}t$! 因为 $t = x^3$, 所以有 $\mathrm{d}t/\mathrm{d}x = 3x^2$. 我们可以把 $\mathrm{d}x$ 移到等式的右侧, 这样有 $\mathrm{d}t = 3x^2\mathrm{d}x$. 先不要考虑这意味着什么, 我们将会在 18.1.3 节中讨论. 好, 现在把等式两端同时除以 3 得 $\frac{1}{3}\mathrm{d}t = x^2\mathrm{d}x$. 这样, 对于原积分函数, 我们可以去掉 x^2 和 $\mathrm{d}x$, 而用 $\frac{1}{3}\mathrm{d}t$ 去替代, 像这样:

$$\int x^2\cos(x^3)\mathrm{d}x = \int \cos(x^3)(x^2\mathrm{d}x) = \int \cos(t)\left(\frac{1}{3}\mathrm{d}t\right).$$

中间的过程不是很必要, 但把 x^2 和 $\mathrm{d}x$ 放到一起可以帮助我们更清楚地看到它们被 $\mathrm{d}t/3$ 所代替. 无论如何, 现在我们都可以把 1/3 移到积分符号的外边, 然后再求积分; 所以有

$$\int x^2\cos(x^3)\mathrm{d}x = \int \cos(t)\frac{1}{3}\mathrm{d}t = \frac{1}{3}\int \cos(t)\mathrm{d}t = \frac{1}{3}\sin(t) + C.$$

如果仅仅把答案写为 $\frac{1}{3}\sin(t) + C$, 这样未免有些懒. 我们从变量 x 开始, 然后变为变量 t, 现在让我们再变回 x. 这样做并不难：仅仅用 x^3 替代 t 即可. 所以最后的答案为

$$\int x^2\cos(x^3)\mathrm{d}x = \frac{1}{3}\sin(x^3) + C.$$

我们可以通过求 $\frac{1}{3}\sin(x^3)$ 对 x 的导数来校验这个结果是否正确.

让我们再来看些例子. 首先, 考虑

$$\int \mathrm{e}^{2x}\sec^2(\mathrm{e}^{2x})\mathrm{d}x.$$

因为 $\sec^2$ 里面的变量是讨厌的 e^{2x}, 所以我们用 t 去替代它. 假设 $t = \mathrm{e}^{2x}$. 求导可得 $\mathrm{d}t/\mathrm{d}x = 2\mathrm{e}^{2x}$. 现在把 $\mathrm{d}x$ 移到右边可得 $\mathrm{d}t = 2\mathrm{e}^{2x}\mathrm{d}x$. 这几乎是积分符号里面的样子, 我们仅仅需要去掉因子 2. 所以两端同时除以 2 可得 $\frac{1}{2}\mathrm{d}t = \mathrm{e}^{2x}\mathrm{d}x$. 把刚才的积分变为以 t 为变量, 我们有

$$\int \mathrm{e}^{2x}\sec^2(\mathrm{e}^{2x})\mathrm{d}x = \int \sec^2(\mathrm{e}^{2x})(\mathrm{e}^{2x}\mathrm{d}x) = \int \sec^2(t)\left(\frac{1}{2}\mathrm{d}t\right).$$

现在把因子 1/2 移出积分符号, 然后积分可得 $\tan(t) + C$. 最后再回到以 x 为变量的状态, 只要用 e^{2x} 替代 t. 这样我们证明了

$$\int \mathrm{e}^{2x}\sec^2(\mathrm{e}^{2x})\mathrm{d}x = \frac{1}{2}\tan(\mathrm{e}^{2x}) + C.$$

再一次提醒, 你可以通过对右边求导来校验这个结果是否正确.

来看另一个例子:

$$\int \frac{3x^2+7}{x^3+7x-9}\mathrm{d}x.$$

这个例子看起来很难. 幸运的是, 如果我们对分母 x^3+7x-9 求导可得 $3x^2+7$. 因此可以通过设 $t=x^3+7x-9$ 来求解. 因为 $\mathrm{d}t/\mathrm{d}x=3x^2+7$, 所以可以写为 $\mathrm{d}t=(3x^2+7)\mathrm{d}x$. 用 t 做变量, 积分的结果为

$$\int \frac{3x^2+7}{x^3+7x-9}\mathrm{d}x=\int \frac{1}{x^3+7x-9}((3x^2+7)\mathrm{d}x)=\int \frac{1}{t}\mathrm{d}t=\ln|t|+C.$$

现在用 x^3+7x-9 来替代 t, 可以回到以 x 为变量的状态. 这样结果为

$$\int \frac{3x^2+7}{x^3+7x-9}\mathrm{d}x=\ln|x^3+7x-9|+C.$$

实际上, 这是一种特殊情况：如果 f 是可导函数, 那么

$$\boxed{\int \frac{f'(x)}{f(x)}\mathrm{d}x=\ln|f(x)|+C.}$$

所以如果分子为分母的导数, 这时的积分结果恰恰是分母的对数(分母要取绝对值并加 C). 我们可以通过设 $t=f(x)$ 来证明. 这时 $\mathrm{d}t/\mathrm{d}x=f'(x)$, 所以有 $\mathrm{d}t=f'(x)\mathrm{d}x$. 如果用链式法则把由 x 为变量变为由 t 为变量的状态, 这时就回到

$$\int \frac{f'(x)}{f(x)}\mathrm{d}x=\int \frac{1}{f(x)}(f'(x)\mathrm{d}x)=\int \frac{1}{t}\mathrm{d}t=\ln|t|+C=\ln|f(x)|+C.$$

这事实上是说, 在上述例子

$$\int \frac{3x^2+7}{x^3+7x-9}\mathrm{d}x$$

中, 可以把答案写为 $\ln|x^3+7x-9|+C$, 因为分子恰恰为分母的导数. 有时分子是分母导数的倍数, 像这样：

$$\int \frac{x}{x^2+8}\mathrm{d}x.$$

分母的导数是 $2x$, 但在分子的位置仅有 x. 没问题, 乘以一个 2 再除以一个 2, 像这样：

$$\int \frac{x}{x^2+8}\mathrm{d}x=\frac{1}{2}\int \frac{2x}{x^2+8}\mathrm{d}x.$$

现在你可以把答案写为 $\dfrac{1}{2}\ln|x^2+8|+C$ 了, 因为分子 $(2x)$ 恰恰是分母 (x^2+8) 的导数. 最后考虑

$$\int \frac{1}{x\ln(x)}\mathrm{d}x.$$

做这道题的最好方法是把这个积分重写为

$$\int \frac{1/x}{\ln(x)} dx,$$

请注意分母 $\ln(x)$ 的导数就是分子 $1/x$. 根据刚才方框里的公式, 这个积分的结果为 $\ln|\ln(x)| + C$, 这样有

$$\int \frac{1}{x\ln(x)} dx = \ln|\ln(x)| + C.$$

18.1.1　换元法和定积分

在定积分中也可以使用换元法. 解决这样的问题有两种方法. 例如, 计算

$$\int_0^{\sqrt[3]{\pi/2}} x^2 \cos(x^3) dx,$$

可以先计算不定积分 $\int x^2 \cos(x^3) dx$, 然后把积分上下限写上. 在上一节中, 我们已经计算过这个不定积分了. 为了简便, 我们用 $t = x^3$ 做换元, 注意 $dt = 3x^2 dx$, 所以 $\frac{1}{3} dt = x^2 dx$, 这时为

$$\int x^2 \cos(x^3) dx = \int \cos(t) \frac{dt}{3} = \frac{1}{3} \int \cos(t) dt = \frac{1}{3} \sin(t) + C = \frac{1}{3} \sin(x^3) + C.$$

事实上最后一步换回到 x 很重要. 无论如何, 关键是我们已经找到了它的导数为 $x^2 \cos(x^3)$, 并且可以使用 17.3 节中微积分的第二基本定理去解决问题：

$$\int_0^{\sqrt[3]{\pi/2}} x^2 \cos(x^3) dx = \frac{1}{3} \sin(x^3) \Big|_0^{\sqrt[3]{\pi/2}} = \left(\frac{1}{3} \sin((\sqrt[3]{\pi/2})^3)\right) - \left(\frac{1}{3} \sin(0^3)\right),$$

通过计算可得结果为 1/3. 所以使用换元法计算定积分的一个方法是：先求不定积分, 然后分别代入积分上下限去求定积分.

这里还有一个方法! 在整个计算过程中你都一直计算定积分, 但一定要记住, 把积分的上下限也用变量 t 来表示. 我们在例子中, 用 $t = x^3$ 去换元, 然后使用 $\frac{1}{3} dt = x^2 dx$ 换到以变量为 t 的积分. 现在当 $x = 0$ 时, 我们有 $t = 0^3 = 0$, 所以积分下限为 0. 但对于积分上限, 当 $x = \sqrt[3]{\pi/2}$ 时, 我们有 $t = \left(\sqrt[3]{\pi/2}\right)^3 = \pi/2$. 这就是说, 我们必须把积分上限换为 $\pi/2$. 综上所述, 这道题换元后的结果为：

$$\int_0^{\sqrt[3]{\pi/2}} x^2 \cos(x^3) dx = \frac{1}{3} \int_0^{\pi/2} \cos(t) dt.$$

我们很快就会完成这道题目, 但请注意如果把积分上限写成这样就大错特错了：

$$\frac{1}{3} \int_0^{\sqrt[3]{\pi/2}} \cos(t) dt,$$

因为我们现在正在对 t 而不是 x 求积分, 因此积分上下限也是以 t 为变量. 事实上, 我们可以以被积函数的变量为积分上下限的变量, 从而使该积分更容易求解, 像这

样：

$$\int_{x=0}^{x=\sqrt[3]{\pi/2}} x^2\cos(x^3)\mathrm{d}x = \frac{1}{3}\int_{t=0}^{t=\pi/2}\cos(t)\mathrm{d}t.$$

这真的强调了我们正在做的事：当 $x=0$ 时, t 也为 0; 当 $x=\sqrt[3]{\pi/2}$ 时, $t=\pi/2$. 所以我们总结一下, 实际上有三次换元：

(1) $\mathrm{d}x$—— 要用 $\mathrm{d}t$ 来表示它, 可以借用被积函数里的其他带有 x 的项来做相应的变化;

(2) 把被积函数里所有带有 x 的项都用 t 来表示;

(3) 积分上下限也用 t 来表示.

让我们来完成这道题目. 最好的方法是把计算过程写在左边, 像这样：

$$t = x^3$$
$$\mathrm{d}t = 3x^2\mathrm{d}x, \text{所以} x^2\mathrm{d}x = \frac{1}{3}\mathrm{d}t$$
$$\text{当} x=0, t=0$$
$$\text{当} x=\sqrt[3]{\pi/2}, t=\pi/2$$

$$\begin{aligned}\int_0^{\sqrt[3]{\pi/2}} x^2\cos(x^3)\mathrm{d}x &= \int_0^{\pi/2}\frac{1}{3}\cos(t)\mathrm{d}t\\ &= \frac{1}{3}\sin(t)\Big|_0^{\pi/2}\\ &= \left(\frac{1}{3}\sin(\pi/2)\right)-\left(\frac{1}{3}\sin(0)\right)=\frac{1}{3}.\end{aligned}$$

注意, 在开始计算右边之前要先把左边的准备工作做完, 因为我们要使用左边所有的信息完成以 t 为变量的转换.

这里还有一个看起来更复杂的例子：

$$\int_{1/\sqrt{2}}^{\sqrt{3}/2}\frac{1}{\sin^{-1}(x)\sqrt{1-x^2}}\mathrm{d}x.$$

先问问你自己：在这个被积函数中, 有哪一项是另一项的导数吗？我们很幸运, $\sin^{-1}(x)$ 的导数为 $1/\sqrt{1-x^2}$. 所以试着做替换 $t=\sin^{-1}(x)$. 的确, $\mathrm{d}t/\mathrm{d}x = 1/\sqrt{1-x^2}$, 所以有

$$\mathrm{d}t = \frac{1}{\sqrt{1-x^2}}\mathrm{d}x.$$

我们还需要把积分上下限用 t 来替代, 把 $x=1/\sqrt{2}$ 和 $x=\sqrt{3}/2$ 分别代入 $t=\sin^{-1}(x)$, 分别得到 $t=\pi/4$ 和 $t=\pi/3$, 只要你还记得反三角函数的基本知识! (参见第 10 章复习这方面的知识.) 我们把这些东西综合在一起, 可得：

$$t=\sin^{-1}(x)$$
$$\mathrm{d}t = \frac{1}{\sqrt{1-x^2}}\mathrm{d}x$$
$$\text{当} x=\frac{1}{\sqrt{2}}, t=\sin^{-1}\left(\frac{1}{\sqrt{2}}\right)=\frac{\pi}{4}$$
$$\text{当} x=\frac{\sqrt{3}}{2}, t=\sin^{-1}\left(\frac{\sqrt{3}}{2}\right)=\frac{\pi}{3}$$

$$\begin{aligned}&\int_{1/\sqrt{2}}^{\sqrt{3}/2}\frac{1}{\sin^{-1}(x)\sqrt{1-x^2}}\mathrm{d}x\\ &= \int_{\pi/4}^{\pi/3}\frac{1}{t}\mathrm{d}t = \ln|t|\Big|_{\pi/4}^{\pi/3}\\ &= \ln\left|\frac{\pi}{3}\right|-\ln\left|\frac{\pi}{4}\right| = \ln\left(\frac{4}{3}\right).\end{aligned}$$

为得到这个最后的化简结果, 我们需要知道基本的对数运算法则 (参见 9.1.4 节). 最好你能记住这些.

顺便说一下, 如果你目光锐利, 会注意到上述换元实际上是上一节最后例子的特例. 这给了我们一个新方法去求解积分

$$\int_{1/\sqrt{2}}^{\sqrt{3}/2} \frac{1}{\sin^{-1}(x)\sqrt{1-x^2}}\mathrm{d}x.$$

让我们从不定积分开始, 把它重写为

$$\int \frac{1}{\sin^{-1}(x)\sqrt{1-x^2}}\mathrm{d}x = \int \frac{1/\sqrt{1-x^2}}{\sin^{-1}(x)}\mathrm{d}x.$$

注意分子正好是分母的导数, 所以我们要取分母的绝对值的对数, 得到最后的答案为

$$\int \frac{1}{\sin^{-1}(x)\sqrt{1-x^2}}\mathrm{d}x = \ln|\sin^{-1}(x)| + C.$$



现在为了计算这个定积分, 可以把原始的积分上下限 $1/\sqrt{2}$ 和 $\sqrt{3}/2$, 一次一个地代入 $\ln|\sin^{-1}(x)|$ 表达式, 然后再求差. 我把计算的细节留给你去做.

这儿有一个关于换元法的不同的问题. 在 16.1.1 节中, 我们说过

如果 f 是一个奇函数, 这时对于任何 a 都有 $\displaystyle\int_{-a}^{a} f(x)\mathrm{d}x = 0.$

你怎样证明这是正确的呢？我们从在 $x=0$ 点把这个积分分成两部分开始：

$$\int_{-a}^{a} f(x)\mathrm{d}x = \int_{-a}^{0} f(x)\mathrm{d}x + \int_{0}^{a} f(x)\mathrm{d}x.$$

对于等式右边的第一个积分, 用 $t=-x$ 去替代. 这时 $\mathrm{d}t = -\mathrm{d}x$. 并且我们看到当 $t=-a$ 时, $x=a$; 当 $t=0$ 时, $x=0$. 所以有

$$\int_{-a}^{0} f(x)\mathrm{d}x = -\int_{a}^{0} f(-t)\mathrm{d}t = \int_{0}^{a} f(-t)\mathrm{d}t.$$

在最后一步, 我们用负号切换积分上下限. 因为函数 f 为奇函数, 所以 $f(-t) = -f(t)$, 这表明了

$$\int_{0}^{a} f(-t)\mathrm{d}t = -\int_{0}^{a} f(t)\mathrm{d}t.$$

现在, 如果我们把虚拟变量变回 x, 那就证明了随后的这个结果：

$$\int_{-a}^{0} f(x)\mathrm{d}x = -\int_{0}^{a} f(x)\mathrm{d}x.$$

这个等式只有在函数 f 为奇函数时才成立! 总之, 我们可以回到最初的方程并使用刚才的这个结果：

$$\int_{-a}^{a} f(x)\mathrm{d}x = \int_{-a}^{0} f(x)\mathrm{d}x + \int_{0}^{a} f(x)\mathrm{d}x = -\int_{0}^{a} f(x)\mathrm{d}x + \int_{0}^{a} f(x)\mathrm{d}x = 0.$$

我们的任务完成了!

18.1.2 如何换元

你怎样选择被替代的函数呢? 这是个很好的问题. 基本思想是寻找其导数也在被积函数中的那些部分. 在积分

$$\int \frac{1}{\sin^{-1}(x)\sqrt{1-x^2}}\mathrm{d}x$$

中, 我们选择 $t=\sin^{-1}(x)$ 换元, 因为它的导数 $1/\sqrt{1-x^2}$ 也恰恰在被积函数中. 在下列积分中这个方法也适用:

$$\int \frac{\sin^{-1}(x)}{\sqrt{1-x^2}}\mathrm{d}x, \quad \int \frac{\mathrm{e}^{\sin^{-1}(x)}}{\sqrt{1-x^2}}\mathrm{d}x \quad 和 \quad \int \frac{1}{\sqrt{\sin^{-1}(x)(1-x^2)}}\mathrm{d}x.$$

在以 t 为变量的情况下, 这些积分分别变为:

$$\int t\mathrm{d}t, \quad \int \mathrm{e}^t\mathrm{d}t \quad 和 \quad \int \frac{1}{\sqrt{t}}\mathrm{d}t,$$

前两个积分很容易观察出来, 但第三个就不那么容易了, 需要把平方根拆开并观察使用换元法怎样计算

$$\int \frac{1}{\sqrt{\sin^{-1}(x)(1-x^2)}}\mathrm{d}x = \int \frac{1}{\sqrt{\sin^{-1}(x)}}\frac{1}{\sqrt{1-x^2}}\mathrm{d}x.$$

现在确定你能准确计算出上述以 t 为变量的积分, 然后可以再把它们替换回以 x 为变量的积分. (对于第三个积分, 如果把 $1/\sqrt{t}$ 写为 $t^{-1/2}$ 会方便我们计算.) 无论怎样计算, 你都会分别得到

$$\frac{(\sin^{-1}(x))^2}{2}+C, \quad \mathrm{e}^{\sin^{-1}(x)}+C \quad 和 \quad 2\sqrt{\sin^{-1}(x)}+C,$$

把得到的每一个结果求导以校验我们的计算是否正确.

有时怎样换元并不很明显. 例如, 怎样计算积分

$$\int \frac{\mathrm{e}^x}{\mathrm{e}^{2x}+1}\mathrm{d}x?$$

可能想象出来的替代为 $t=\mathrm{e}^x, t=\mathrm{e}^{2x}, t=\mathrm{e}^{2x}+1$. 后两个换元并不能解决问题, 因为在这两种情况中 $\mathrm{d}t=2\mathrm{e}^{2x}\mathrm{d}x$, 但被积函数的分子中没有 e^{2x} 这项. 所以我们设 $t=\mathrm{e}^x$, 这时有 $\mathrm{d}t=\mathrm{e}^x\mathrm{d}x$, 该项正好出现在分子. 至于分母, 我们可以把 e^{2x} 改写为 $(\mathrm{e}^x)^2$, 这就是 t^2. 所以

$$\int \frac{\mathrm{e}^x}{\mathrm{e}^{2x}+1}\mathrm{d}x = \int \frac{1}{t^2+1}\mathrm{d}t,$$

这个积分的结果为 $\tan^{-1}(t)+C$. 再把它换回以 x 为变量的积分, 有

$$\int \frac{\mathrm{e}^x}{\mathrm{e}^{2x}+1}\mathrm{d}x = \tan^{-1}(\mathrm{e}^x)+C$$


对于任何常数 C 都成立. 我们可以对等式的右端求导来校验这个结果是否正确.

我们再看一个例子：

$$\int x\sqrt[5]{3x+2}\mathrm{d}x.$$

关于计算 $\sqrt[n]{ax+b}$ 这种类型的积分, 我们有一个非常好的方法. 可以简单地设 $t=\sqrt[n]{ax+b}$, 在求 $\mathrm{d}t$ 之前先两边同时 n 次方. 所以：

在换掉 $\sqrt[n]{ax+b}$ 之前, 设 $t=\sqrt[n]{ax+b}$ 并对等式 $t^n=ax+b$ 两端求导.

所以, 我们设 $t=\sqrt[5]{3x+2}$. 为找到 $\mathrm{d}t$, 把等式两端 5 次方, 得 $t^5=3x+2$. 现在把这个新的等式两端同时对合适的变量 (由链式法则决定) 求导, 得 $5t^4\mathrm{d}t=3\mathrm{d}x$. $5t^4$ 是 t^5 关于 t 的导数, 3 是 $3x$+2 关于 x 的导数. 所以, 我们找到了可以用 t 表示 3 $\mathrm{d}x$ 的表达式, 等式两端同时除以 3 就得到了 $\mathrm{d}x$ 的表达式. 在本例中有

$$\mathrm{d}x=\frac{5}{3}t^4\mathrm{d}t.$$

(通过写出关于 x 的表达式 $x=\dfrac{1}{3}\left(t^5-2\right)$, 再两边同时对 t 求导也可得上式.) 现在让我们重新看看这个积分. 该积分表达式有三项：x、$\sqrt[5]{3x+2}$ 和 $\mathrm{d}x$. 第二项就是 t, 我们已经找到了用 t 表示第三项的表达式. 那么第一项 x 该怎么处理呢? 我们知道 $t^5=3x+2$, 所以可以重新整理这个等式得 $x=\dfrac{1}{3}\left(t^5-2\right)$. 这样, 这个积分表达式为

$$\int x\sqrt[5]{3x+2}\mathrm{d}x=\int\frac{1}{3}(t^5-2)(t)\times\frac{5}{3}t^4\mathrm{d}t.$$

现在我们先做乘法然后再积分可得

$$\frac{5}{9}\int(t^{10}-2t^5)\mathrm{d}t=\frac{5}{99}t^{11}-\frac{5}{27}t^6+C.$$

再回到以 x 为变量的积分：用 $t=(3x+2)^{1/5}$ 再替换得

$$\frac{5}{99}(3x+2)^{11/5}-\frac{5}{27}(3x+2)^{6/5}+C.$$


你应该试着自己解决这个问题, 把你的解题思路写在计算过程的左边, 就如前面例子中演示过的那样. 并且你也应该把所得的积分结果求导, 校验是否会得到 $x\sqrt[5]{3x+2}$. 顺便说一下, 你是否注意到这道题用的换元法同我们以前的换元有何不同? 确实有一点点不同. 在其他例子中, 我们用的方程是 $\mathrm{d}t$=(关于 x 的函数)$\mathrm{d}x$, 然而在这个例子中我们写为 $\mathrm{d}x=\dfrac{5}{3}t^4\mathrm{d}t$. 这种写法对于计算很有帮助, 因为我们可以直接替代 $\mathrm{d}x$. 在所有的其他例子中, 我们不得不找到一个已经存在的 x 的表达式的常数倍才有机会化简. 在 19.3 节中, 我们将要看到能直接替代 $\mathrm{d}x$ 的其他例子.

总的来说, 对于怎样换元没有硬性规定. 你需要跟着直觉走, 这个直觉只有在你做了大量的习题之后才会越来越正确. 可以尝试任何你想到的换元. 如果换元后的积分比原始的积分更糟糕, 或者你找不到任何方法把每一个变量都化成 t 变量, 那也不要着急：你仅仅需要做的是再回到原始积分, 然后尝试其他的换元.



现在, 在介绍分部积分法之前我需要再阐述两点. 第一点是换元方法的识别, 我将在下一节介绍这点. 第二点是对换元法的总结, 即

- 对于**不定积分**, 用 t 和 $\mathrm{d}t$ 分别表示带有 x 的表达式和 $\mathrm{d}x$, 然后再求这个新的用 t 表达的积分, 最后再换回到 x;
- 对于**定积分**, 用 t 和 $\mathrm{d}t$ 分别表示带有 x 的表达式和 $\mathrm{d}x$, **并且**也要把积分上下限换为与 t 相关, 这时计算这个新的积分 (没有必要再回到以 x 为变量的状态). 当然也可以用另一个方法, 就是先把它看成不定积分去计算结果, 然后再把积分上下限分别代入求最后的结果.

18.1.3 换元法的理论解释



假设你想在某些积分中做这样的换元 $t = x^2$, 这样就得到 $\mathrm{d}t/\mathrm{d}x = 2x$, 改写为 $\mathrm{d}t = 2x\mathrm{d}x$. 在某种意义下, 这是一种没有意义的陈述 —— 毕竟 $\mathrm{d}t$ 和 $\mathrm{d}x$ 没什么实际意义. 我们知道 $\mathrm{d}t/\mathrm{d}x$ 是导数的一种表示, 但在第 13 章中 $\mathrm{d}t$ 和 $\mathrm{d}x$ 仅仅被定义为微分. 所以 $\mathrm{d}t = 2x\mathrm{d}x$ 究竟意味着什么？一个好的解释是, 当 t 发生微小变化时, 它的变化量是它所对应的 x 的微小变化量的 $2x$ 倍. 实际上在 5.2.7 节中有过这种类型的表达式. 你可以采用这种方式去观察它, 看怎样用黎曼和解释它, 但这里有一个更好的方式：仅仅使用链式法则.

设想你已经做了一个换元 $t = g(x)$, 我们用以 t 为变量的 $\int f(t)\mathrm{d}t$ 结束求解, 设结果为 $F(t) + C(C$ 为常数). 所以这个积分以 t 为变量可写为

$$\int f(t)\mathrm{d}t = F(t) + C.$$

因为 $t = g(x)$, 所以可得 $\mathrm{d}t = g'(x)\mathrm{d}x$, 这样上述方程可以转化为以 x 为变量的式子：

$$\int f(g(x))g'(x)\mathrm{d}x = F(g(x)) + C.$$

我所做的就是分别用 $g(x)$ 替代 t, 用 $g'(x)\mathrm{d}x$ 替代 $\mathrm{d}t$. 如果你想证明这个替代是有效的, 我们需要证明上述等式是正确的. 设 $h(x) = F(g(x))$, 根据链式法则 (参见 6.2.5 节第一部分), $h'(x) = F'(g(x))g'(x)$ 是正确的. 我们可以以不定积分的形式来表达：

$$\int F'(g(x))g'(x)\mathrm{d}x = h(x) + C.$$

因为 $h(x) = F(g(x))$, 我们有

$$\int F'(g(x))g'(x)\mathrm{d}x = F(g(x)) + C.$$

现在, 因为 $\int f(t)\mathrm{d}t = F(t) + C$, 所以 $F'(t) = f(t)$; 因为 $t = g(x)$, 我们有 $F'(g(x)) = f(g(x))$. 这样上述等式变为

$$\int f(g(x))g'(x)\mathrm{d}x = F(g(x)) + C.$$

这正是我们要证明的!

顺便说一下, 这个漂亮的等式可以帮助我们证明一种换元法, 这种方法恰恰就是我们在上一节最后例子之后讨论过的. (当我们学习 19.3 节的三角换元法时, 将会一次又一次地见到这个方法). 第二种换元法是, 不设 $t = g(x)$ 而是对于一些函数 g 设 $x = g(t)$, 这样用 $g'(t)\mathrm{d}t$ 替代 $\mathrm{d}x$. 在这种情况下, 最初的积分 $\int f(x)\mathrm{d}x$ 现在变为

$$\int f(g(t))g'(t)\mathrm{d}t.$$

现在可以计算出这个积分了, 然后再回到以 x 为变量的积分上. 根据我们刚才证明的漂亮的等式, 其中用 t 替代了 x, 我们看到上述积分等于 $F(g(t)) + C$, 其中 F 是 f 的反导数. 这时的结果恰恰就是 $F(x) + C$, 这正是我们想要的. 所以这个方法很有用, 我们证明了此换元法的合理性.

18.2　分部积分法

我们已经看到换元法是怎样逆用链式求导法则的. 还有一种方法可以逆用乘积法则, 我们称之为分部积分法. 让我们回忆一下 6.2.3 节的乘积法则：如果 u 和 v 是关于 x 的函数, 则有

$$\frac{\mathrm{d}}{\mathrm{d}x}(uv) = v\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}v}{\mathrm{d}x}.$$

让我们重新写一下这个等式, 然后两边同时再对 x 求积分, 得到

$$\int u\frac{\mathrm{d}v}{\mathrm{d}x}\mathrm{d}x = \int \frac{\mathrm{d}}{\mathrm{d}x}(uv)\mathrm{d}x - \int v\frac{\mathrm{d}u}{\mathrm{d}x}\mathrm{d}x.$$

等式右侧的第一项是函数uv导数的反导数, 所以它等于uv+C. 其实 $+C$ 是不必要的, 因为等式右侧的第二项已然是个不定积分：它自动包含一个 $+C$. 所以我们已经证明了

$$\int u\frac{\mathrm{d}v}{\mathrm{d}x}\mathrm{d}x = uv - \int v\frac{\mathrm{d}u}{\mathrm{d}x}\mathrm{d}x.$$

这就是分部积分公式, 这种形式非常实用, 但我们还有这种形式的简单写法, 更方

便. 如果我们用 $\mathrm{d}v$ 替代 $\frac{\mathrm{d}v}{\mathrm{d}x}\mathrm{d}x$, 用 $\mathrm{d}u$ 替代 $\frac{\mathrm{d}u}{\mathrm{d}x}\mathrm{d}x$, 会得到公式

$$\boxed{\int u\mathrm{d}v = uv - \int v\mathrm{d}u.}$$

再一次提醒, 这仅仅是公式的简写形式, 但这种写法确实很实用. 让我们看看它怎样帮助我们解决问题. 假设我们想求解



$$\int x\mathrm{e}^x\mathrm{d}x.$$

换元法看起来不管用了 (试试看是否能解决问题), 所以我们尝试使用分部积分法. 首先要得到 $\int u\mathrm{d}v$ 形式的积分, 这样才能应用分部积分法. 有很多种方法可以化成这种形式, 但有一种很管用的方法：设 $u = x$ 并且 $\mathrm{d}v = \mathrm{e}^x\mathrm{d}x$. 这时我们有 $\int x\mathrm{e}^x\mathrm{d}x = \int u\mathrm{d}v$.

现在我们使用分部积分法, 需要找到 $\mathrm{d}u$ 和 v. $\mathrm{d}u$ 很容易找到：我们知道 $u = x$, 所以 $\mathrm{d}u$=$\mathrm{d}x$. 那么 v 怎样找呢？我们有 $\mathrm{d}v = \mathrm{e}^x\mathrm{d}x$, 所以 v 究竟是多少呢？仅仅对这个等式两侧同时求积分：$\int \mathrm{d}v = \int \mathrm{e}^x\mathrm{d}x$. 这就是说 $v = \mathrm{e}^x + C$. 实际上我们并不需要这样的 v, 仅仅需要能给出 $\mathrm{d}v = \mathrm{e}^x\mathrm{d}x$ 这种形式的 v. 所以我们可以忽略 $+C$ 仅仅设 $v = \mathrm{e}^x$.

我们现在要开始应用分部积分公式了, 其中 $u = x$, $\mathrm{d}u = \mathrm{d}x$, $v = \mathrm{e}^x$, 并且 $\mathrm{d}v = \mathrm{e}^x\mathrm{d}x$. 使用这个公式的最简单方式是留有一定间隔地写下这个公式, 然后进行如下替代：

$$\begin{aligned} \int u \quad \mathrm{d}v \quad &= u\ v \quad - \int \ v \ \ \mathrm{d}u \\ \int x\ \overbrace{\mathrm{e}^x\mathrm{d}x} &= x\ \mathrm{e}^x - \int \mathrm{e}^x\ \mathrm{d}x. \end{aligned}$$

现在仍然有一个积分被剩下了, 唯一被剩下的 $\int \mathrm{e}^x\mathrm{d}x$ 的结果为 $\mathrm{e}^x + C$. 把这个加进去, 就得到 $\int x\mathrm{e}^x\mathrm{d}x = x\mathrm{e}^x - \mathrm{e}^x + C$. (从技术角度来说应该是 $-C$, 而不是 $+C$; 但减一个常数也就是加这个常数的相反数, 区分这个是没有必要的.)

为了能计算出 $\mathrm{d}u$ 和 v, 我建议你这样来写：

$$\begin{aligned} u &= x & v &= \\ \mathrm{d}u &= & \mathrm{d}v &= \mathrm{e}^x\mathrm{d}x, \end{aligned}$$

这时通过对 u 求导和对 $\mathrm{d}v$ 求积分来填写空白处：

$$\begin{aligned} u &= x & v &= \mathrm{e}^x \\ \mathrm{d}u &= \mathrm{d}x & \mathrm{d}v &= \mathrm{e}^x\mathrm{d}x. \end{aligned}$$

你能很容易地用分部表达式替代这个积分, 因为我们已经做好了所有的准备工作.

你究竟为什么决定选择 $u = x$ 和 $\mathrm{d}v = \mathrm{e}^x\mathrm{d}x$ 呢？ 为什么我们不设 $u = \mathrm{e}^x$ 和

$\mathrm{d}v = x\mathrm{d}x$ 呢? 我们可以这样做的. 在这种情况下, 会有

$$\begin{aligned} u &= \mathrm{e}^x & v &= \frac{1}{2}x^2 \\ \mathrm{d}u &= \mathrm{e}^x\mathrm{d}x & \mathrm{d}v &= x\mathrm{d}x; \end{aligned}$$

注意我们通过对 $\mathrm{d}v = x\mathrm{d}x$ 求积分得到 $v = \dfrac{1}{2}x^2$ (记住我们不需要 $+C$). 这时, 通过分部积分法有

$$\begin{aligned} \int u \quad \mathrm{d}v \quad &= u \quad v \quad - \int \quad v \quad \mathrm{d}u \\ \int x\mathrm{e}^x\mathrm{d}x = \int \mathrm{e}^x \overbrace{x\mathrm{d}x} &= \mathrm{e}^x \cdot \frac{1}{2}x^2 - \int \frac{1}{2}x^2 \overbrace{\mathrm{e}^x\mathrm{d}x}. \end{aligned}$$

整个解题过程没有任何错误, 但它很不实用. 你看, 最后这个积分是比原始积分更复杂的积分! 所以我们最好使用第一种方法. 通常来说, 如果你在表达式里面见到 e^x, 好好待它, 它是你的朋友, 因为它的积分是它自己. 这里的规则是：如果 e^x 存在, 通常让 $\mathrm{d}v = \mathrm{e}^x\mathrm{d}x$, 因为这样可以很简单地得到 v 即为 e^x.

一些变形

这里面有很多复杂的情况. 有时你需要多次计算分部积分. 例如, 你怎样计算

$$\int x^2 \sin(x)\mathrm{d}x?$$

很好, 它是一个乘积的形式, 所以换元法不适用, 我们试着用分部积分法. 这里没有 e^x, 但是有 $\sin(x)$, 这也非常好. 让我们设 $u = x^2$, 并且 $\mathrm{d}v = \sin(x)\mathrm{d}x$. 我们得到

$$\begin{aligned} u &= x^2 & v &= -\cos(x) \\ \mathrm{d}u &= 2x\mathrm{d}x & \mathrm{d}v &= \sin(x)\mathrm{d}x; \end{aligned}$$

这里我们通过对 $\mathrm{d}v = \sin(x)\mathrm{d}x$ 求积分可得 $v = \int \sin(x)\mathrm{d}x = -\cos(x)$(记住没有必要写 $+C$). 所以我们有

$$\begin{aligned} \int u \quad \mathrm{d}v \quad &= u \quad v \quad - \int \quad v \quad \mathrm{d}u \\ \int x^2 \overbrace{\sin(x)\mathrm{d}x} &= x^2 \overbrace{(-\cos(x))} - \int \overbrace{(-\cos(x))}\,\overbrace{2x\mathrm{d}x} \\ &= -x^2\cos(x) + \int \cos(x) \cdot 2x\mathrm{d}x. \end{aligned}$$

现在我们把 2 从最后的积分中提出来, 我们将要完成这个积分, 只要知道 $\int x\cos x\mathrm{d}x$ 的积分结果. 这比我们的原始积分表达式要简单, 原始表达式里是 x^2 而现在仅仅为 x 了, 毕竟余弦函数和正弦函数是非常相似的. 所以再一次用分部积分法. 我们假设 $U = x$ 并且 $\mathrm{d}V = \cos(x)\mathrm{d}x$; 这次我使用大写字母是因为前面已使用了小写字母. 现在我们有

$$\begin{aligned} U &= x & V &= \sin(x) \\ \mathrm{d}U &= \mathrm{d}x & \mathrm{d}V &= \cos(x)\mathrm{d}x, \end{aligned}$$

这样, 通过替代有

$$\int U \quad \mathrm{d}V \quad = U \quad V \quad - \int \quad V \quad \mathrm{d}U$$
$$\int x \overbrace{\cos(x)\mathrm{d}x} = x \ \sin(x) - \int \sin(x) \ \mathrm{d}x.$$

我们已经知道 $\int \sin(x)\mathrm{d}x = -\cos(x) + C$, 所以有

$$\int x\cos(x)\mathrm{d}x = x\sin(x) + \cos(x) + C.$$

我们几乎快完成了, 仅仅需要做的是把这些代入最开始的表达式里：

$$\int x^2 \sin(x)\mathrm{d}x = -x^2\cos(x) + 2x\sin(x) + 2\cos(x) + C.$$

(再一次强调, 我不写 $+2C$ 因为它只是一个常数.)

有时在两次分部积分之后情况并未好转. 在这种情况下, 如果你运气好, 那么将会得到原始积分的倍数. 如果你很不走运, 那只有把刚才的计算扔到一边, 重新再来了. (如果你很不走运, 那么原始积分就可能会被正好约掉, 这样就一点忙也帮不上了!) 这种情况到底是什么样的, 这里有一个例子：



$$\int \cos(x)\mathrm{e}^{2x}\mathrm{d}x.$$

这个被积函数既包含余弦函数又包含指数函数, 但我更倾向于用指数函数, 所以, 我们设 $u = \cos(x), \mathrm{d}v = \mathrm{e}^{2x}\mathrm{d}x$. 得到

$$\begin{aligned} u &= \cos(x) & v &= \frac{1}{2}\mathrm{e}^{2x} \\ \mathrm{d}u &= -\sin(x)\mathrm{d}x & \mathrm{d}v &= \mathrm{e}^{2x}\mathrm{d}x. \end{aligned}$$

(当你对 e^{2x} 求积分以求 v 时, 别忘记要除以 2.) 这样, 我们有

$$\int \quad u \quad \mathrm{d}v \quad = \quad u \quad v \quad - \int \quad v \quad \mathrm{d}u$$
$$\int \cos(x) \overbrace{\mathrm{e}^{2x}\mathrm{d}x} = \cos(x) \overbrace{\frac{1}{2}\mathrm{e}^{2x}} - \int \overbrace{\frac{1}{2}\mathrm{e}^{2x}} \overbrace{(-\sin(x))\mathrm{d}x}$$
$$= \frac{1}{2}\cos(x)\mathrm{e}^{2x} + \frac{1}{2}\int \sin(x)\mathrm{e}^{2x}\mathrm{d}x.$$

现在等式右侧的新积分同我们最开始计算的积分表达式的难度是等同的, 所以我们选择的这种计算方法是否合适还不是很清楚. 无论如何, 我们先坚持这种方法, 再次用分部积分法. 这次我们设 $U = \sin(x), \mathrm{d}V = \mathrm{e}^{2x}\mathrm{d}x$. 我们看看得到什么：

$$\begin{aligned} U &= \sin(x) & V &= \frac{1}{2}\mathrm{e}^{2x} \\ \mathrm{d}U &= \cos(x)\mathrm{d}x & \mathrm{d}V &= \mathrm{e}^{2x}\mathrm{d}x. \end{aligned}$$

通过分部积分法, 我们有

$$\int \quad U \quad dV = U \quad V - \int \quad V \quad dU$$

$$\int \sin(x) \overbrace{e^{2x}dx} = \sin(x) \overbrace{\frac{1}{2}e^{2x}} - \int \overbrace{\frac{1}{2}e^{2x}} \overbrace{\cos(x)dx}$$

$$= \frac{1}{2}\sin(x)e^{2x} - \frac{1}{2}\int \cos(x)e^{2x}dx.$$

把这两次计算合并到一起, 有

$$\int \cos(x)e^{2x}dx = \frac{1}{2}\cos(x)e^{2x} + \frac{1}{2}\left(\frac{1}{2}\sin(x)e^{2x} - \frac{1}{2}\int \cos(x)e^{2x}dx\right)$$

$$= \frac{1}{2}\cos(x)e^{2x} + \frac{1}{4}\sin(x)e^{2x} - \frac{1}{4}\int \cos(x)e^{2x}dx.$$

这能帮助我们计算吗？是的. 我们注意到等式两端出现了同样的积分, 再把这两个积分都移到等式的左边. 事实上, 我们可以在等式两侧同时加上原始积分的 1/4, 这样可以把等式右边的积分消掉, 再加上一个常数 C 得到

$$\frac{5}{4}\int \cos(x)e^{2x}dx = \frac{1}{2}\cos(x)e^{2x} + \frac{1}{4}\sin(x)e^{2x} + C.$$

现在我们在等式两侧同时乘以 4/5 得到

$$\int \cos(x)e^{2x}dx = \frac{2}{5}\cos(x)e^{2x} + \frac{1}{5}\sin(x)e^{2x} + C.$$

(再一次提醒, 我们不写 $+\frac{4}{5}C$, 仅仅写 $+C$ 来表示常数.)

这里有另一种类型的积分也需要用到分部积分法, 但是它的计算更复杂. 在这种情况下, 这种类型的积分没有乘积的形式. 这种类型的积分有：

$$\int \ln(x)dx, \quad \int (\ln(x))^2dx, \quad \int \sin^{-1}(x)dx, \quad \int \tan^{-1}(x)dx.$$

这就是说, 如果积分是反三角函数或 $\ln(x)$ 的幂的形式, 可以用分部积分法. 在这种情况下, 应该设 u 为这个函数本身, 并让 $dv = dx$. 例如, 计算

$$\int_0^1 \tan^{-1}(x)dx,$$

我们设 $u = \tan^{-1}(x), dv = dx$, 这时有

$$u = \tan^{-1}(x) \qquad v = x$$

$$du = \frac{1}{1+x^2}dx \qquad dv = dx,$$

并且有 (我们暂时忽略积分上下限)

$$\int u dv = uv - \int v du$$

$$\int \tan^{-1}(x)dx = \tan^{-1}(x)x - \int x\frac{1}{1+x^2}dx$$

$$= x\tan^{-1}(x) - \int \frac{x}{1+x^2}dx.$$

使用 18.1 节最后的方法, 右侧的积分等于 $\frac{1}{2}\ln(1+x^2)+C$(确信你同意这点!), 所以我们有



$$\int_0^1 \tan^{-1}(x)\mathrm{d}x = \left(x\tan^{-1}(x) - \frac{1}{2}\ln(1+x^2)\right)\Big|_0^1 = \frac{\pi}{4} - \frac{1}{2}\ln(2).$$

你怎么得到这个最后答案的呢? 我们知道对数和反三角函数! 确保你相信上述答案是正确的. 同时注意, 我们先计算不定积分以便计算定积分 (我们要先把积分变量转移到以 u 和 v 为变量上来). 这通常是一个很好的方法. 也就是说, 当用分部积分法求解一个定积分的表达式时, 先寻找它的不定积分, 最后再把积分上下限代入.





18.3 部分分式

让我们研究怎样对一个有理函数求积分. 我们将要计算积分

$$\int \frac{p(x)}{q(x)}\mathrm{d}x,$$

其中 p 和 q 为多项式. 这种类型在积分中所占的比例很大, 例如

$$\int \frac{x^2+9}{x^4-1}\mathrm{d}x, \quad \int \frac{x}{x^3+1}\mathrm{d}x \quad 或 \quad \int \frac{1}{x^3-2x^2+3x-7}\mathrm{d}x.$$

这些题目看起来有些复杂. 也有一些简单的例子：

$$\int \frac{1}{x-3}\mathrm{d}x, \quad \int \frac{1}{(x+5)^2}\mathrm{d}x, \quad \int \frac{1}{x^2+9}\mathrm{d}x \quad 和 \quad \int \frac{3x}{x^2+9}\mathrm{d}x.$$

后面这四个积分也是有理函数的积分, 但相对简单一些. 我们可以尽量使用换元法求解这些积分. (这四道题的换元分别是 $t=x-3, t=x+5, t=x/3$ 和 $t=x^2+9$.) 前面两个积分的分母是线性方程的幂, 而后两个是二次函数且不能因式分解.



基本方法：首先会看怎样处理有理函数, 我们通过一些代数运算把它分解成几个更简单的有理函数的和的形式; 然后将看到如何对这些简单的有理函数求积分. 我们提到的这些更简单的有理函数就像刚才最后那四个一样：它们看起来要么像一个常数除以一个线性函数, 要么像一个线性函数除以一个二次函数. 我们将首先研究这些代数运算, 然后再用微积分的方法. 最后我将总结这种方法, 并给出一个完整的例子.

18.3.1 部分分式的代数运算

我们的目的是把一个有理函数分成许多更简单的部分. 第一步是确保这个函数的分子的次数小于分母的次数. 如果不是这样, 我们需要先做一个多项式的除法. 所以在下面这些例子中：

$$\int \frac{x+2}{x^2-1}\mathrm{d}x \quad 和 \quad \int \frac{5x^2+x-3}{x^2-1}\mathrm{d}x,$$

第一例很容易, 因为很显然分子的次数小于分母的次数. 但第二个例子就不那么容易了, 因为分子的次数同分母的次数是一样的 (都是 2). 如果分子是三次或更高次, 那我们会有同样的麻烦. 这样的话, 我们要做一个多项式的除法. 为此, 将其写为

$$分母\overline{)分子}.$$



来看我们的例子

$$\int \frac{5x^2+x-3}{x^2-1}\mathrm{d}x,$$

这道题的除法算式如下：

$$\begin{array}{r} 5 \\ x^2-1\overline{)5x^2+x-3} \\ \underline{5x^2\qquad -5} \\ x+2 \end{array}$$

这个除法告诉我们, 商为 5, 余数为 $x+2$. 所以有

$$\frac{5x^2+x-3}{x^2-1} = 5 + \frac{x+2}{x^2-1}.$$

如果等式两边同时对 x 求积分, 我们有

$$\int \frac{5x^2+x-3}{x^2-1}\mathrm{d}x = \int \left(5 + \frac{x+2}{x^2-1}\right)\mathrm{d}x.$$

现在, 我们能把这个积分分成了两部分, 并且对第一部分求积分, 我们看到原始的积分变为

$$\int 5\mathrm{d}x + \int \frac{x+2}{x^2-1}\mathrm{d}x = 5x + \int \frac{x+2}{x^2-1}\mathrm{d}x.$$

这个新积分的分子次数为 1, 分母次数为 2, 这正是我们想要的结果. 现在我们可以继续计算了.

接下来, 我们将要分解分母中的因式. 如果分母是一个二次函数, 请检验它的判别式：像在 1.6 节中那样, 如果判别式为负, 你不能对它进行因式分解. 否则, 可以手动进行因式分解或使用二次公式. 如果分母很复杂, 你就不得不猜想一个根, 然后再用多项式的除法.



在因式分解分母之后, 下一步要写下的东西叫作 “分部”. 这是通过把分母的一个或更多的因式加到一起构成的, 它依据如下的规则.

(1) 如果有线性式 $(x+a)$, 那么这个分部有如下形式：

$$\frac{A}{x+a}.$$

(2) 如果有线性式的平方 $(x+a)^2$, 那么这个分部有如下形式：

$$\frac{A}{(x+a)^2} + \frac{B}{x+a}.$$

(3) 如果有二次多项式 (x^2+ax+b), 那么这个分部有如下形式：

$$\frac{Ax+B}{x^2+ax+b}.$$

这些都是很常见的例子. 也有一些不常见的例子.

(4) 如果有线性式的三次方 $(x+a)^3$, 那么这个分部有如下形式：

$$\frac{A}{(x+a)^3}+\frac{B}{(x+a)^2}+\frac{C}{x+a}.$$

(5) 如果有线性式的四次方 $(x+a)^4$, 那么这个分部有如下形式：

$$\frac{A}{(x+a)^4}+\frac{B}{(x+a)^3}+\frac{C}{(x+a)^2}+\frac{D}{x+a}.$$

注意: **分部仅仅由分母决定**, 同分子没有关系! 并且当我们使用常数 A、B、C、D 时, 记住不能在不同的表达式里重复使用这些字母. 所以, 你需要使用字母表中后续的字母. 在刚才的例子

$$\int\frac{x+2}{x^2-1}\mathrm{d}x$$

中, 分母是 $(x+1)(x-1)$; 所以我们有两个线性因子, 分部为：

$$\frac{A}{x-1}+\frac{B}{x+1}.$$

我们不能两次使用 A, 所以在第二项使用 B. 顺便说一下, 如果你使用字母 C_1、C_2、C_3 或其他字母, 而不是使用 A、B、C 等字母, 那么是在玩火自焚. 如果这样做, 那你会经常犯一些不经意的错误, 除非你能注意到这些字母下脚标的不同.

这里还有另一个例子.

$$\frac{\text{任意陈旧形式}}{(x-1)(x+4)^3(x^2+4x+7)(3x^2-x+1)}$$

这种形式会怎样呢? 答案是

$$\frac{A}{x-1}+\frac{B}{(x+4)^3}+\frac{C}{(x+4)^2}+\frac{D}{x+4}+\frac{Ex+F}{x^2+4x+7}+\frac{Gx+H}{3x^2-x+1}.$$

你可能会以不同的顺序写下这些项, 或交换从 A 到 H 这些字母的顺序; 这都可以.

一旦写下这种形式, 你也应该写下被积函数等于这种形式, 然后再使用等式两端同时乘以分母的方法. 例如, 我们发现积分

$$\int\frac{x+2}{x^2-1}\mathrm{d}x$$

的被积函数可以写为

$$\frac{A}{x-1}+\frac{B}{x+1};$$

所以我们有

$$\frac{x+2}{x^2-1}=\frac{A}{x-1}+\frac{B}{x+1}.$$

实际上, 最好把等式左侧的分式写为

$$\frac{x+2}{(x-1)(x+1)}=\frac{A}{x-1}+\frac{B}{x+1}.$$

现在等式两侧同时乘以分母可得

$$x+2=A(x+1)+B(x-1).$$

注意, 等式右侧的第一项消掉了 $(x-1)$ 因子, 第二项削掉了 $(x+1)$ 因子. 无论如何, 现在我们有两个方法可以继续. 第一个方法是替代 x. 如果设 $x=1$, 这时 $B(x-1)$ 这项就消失了, 这时我们有

$$1+2=A(1+1).$$

也就是说 $A=\frac{3}{2}$. 现在如果你把 $x=-1$ 代入原始方程, 那么 $A(x+1)$ 这项将会消失：

$$-1+2=B(-1-1).$$

所以 $B=-\frac{1}{2}$. 另一个求解 A 和 B 的方法, 是将原始方程 $x+2=A(x+1)+B(x-1)$, 重写为

$$x+2=(A+B)x+(A-B).$$

现在通过 x 的系数相等可得 $1=A+B$, 通过常数项相等可得 $2=A-B$. 同时解这两个方程, 可得 $A=\frac{3}{2},B=-\frac{1}{2}$, 同刚才的答案是一样的.

你可能已经注意到, 计算 A 和 B 的两种方法都需要两个条件：对于替代法, 把 $x=1$ 和 $x=-1$ 分别代入; 对于系数相等的方法, 让 x 的系数相等, 也让常数项相等. 我们可以选择其中的任意一种方法. 例如, 如果代入 $x=1$ 会得到 $A=\frac{3}{2}$; 如果取 x 的系数相等, 会发现 $1=A+B$, 所以 $B=-\frac{1}{2}$. 在通常情况下, 有多少个常数需要求出, 就需要应用多少次刚才的某种方法, 或者你也可以混着使用两种方法.

我们剩下的工作是重写积分表达式 (用分部的形式), 但这次需要把常数代进去. 所以我们的例子为

$$\frac{x+2}{x^2-1}=\frac{A}{x-1}+\frac{B}{x+1}=\frac{3/2}{x-1}+\frac{-1/2}{x+1}.$$

现在对等式两端同时积分, 并把常数项移到等式的外边, 可得

$$\int\frac{x+2}{x^2-1}\mathrm{d}x=\frac{3}{2}\int\frac{1}{x-1}\mathrm{d}x-\frac{1}{2}\int\frac{1}{x+1}\mathrm{d}x.$$

我们已经成功地把原始复杂的积分化成了两个很简单的积分. 我们会很快求解这两个积分的.

到目前为止, 我们看到除非分子的次数小于分母的次数, 否则都要做一个多项式的除法, 然后对分母进行因式分解, 写下它的每一个分部, 再使用上述两种方法中的一个计算这些未知的常数. 最后我们写下这些积分的各个部分. 我们将在 18.3.3

节中看到另一个解决这类问题的例子. 与此同时, 让我们做一些积分.

18.3.2 对每一部分积分

我们需要知道在把原始积分分成小部分后怎样求解每一部分的积分. 简单的积分形式是

$$\int \frac{1}{ax+b}\mathrm{d}x.$$

为此, 我们可以设 $t = ax + b$. 例如在上一节的最后, 我们看到

$$\int \frac{x+2}{x^2-1}\mathrm{d}x = \frac{3}{2}\int \frac{1}{x-1}\mathrm{d}x - \frac{1}{2}\int \frac{1}{x+1}\mathrm{d}x.$$

对于第一个积分表达式, 可以通过设 $t = x - 1$ 求解, 而第二个可设 $t = x + 1$. 在这两种情况下 $\mathrm{d}t$=$\mathrm{d}x$, 所以很容易得到

$$\int \frac{x+2}{x^2-1}\mathrm{d}x = \frac{3}{2}\ln|x-1| - \frac{1}{2}\ln|x+1| + C.$$

还有另一个例子, 求解

$$\int \frac{1}{4x+5}\mathrm{d}x.$$

我们设 $t = 4x + 5$ 所以有 $\mathrm{d}t$=$4\mathrm{d}x$; 这时积分转移到了以 t 为变量的状态, 变为 $\frac{1}{4}\int 1/t\mathrm{d}t$, 它的结果为 $\frac{1}{4}\ln|t| + C$. 最后再用 x 替代 t, 上述积分结果为 $\frac{1}{4}\ln|4x + 5| + C$.

若分母是线性因式的幂, 这种方法也是适用的. 例如, 计算

$$\int \frac{1}{(4x+5)^2}\mathrm{d}x,$$

可以再一次用 $t = 4x + 5$ 换元. 这样, 这个积分变为 $\frac{1}{4}\int 1/t^2\mathrm{d}t$, 它的积分结果为 $-\frac{1}{4}(1/t) + C$; 再换回到以 x 为变量的状态, 这样就证明了

$$\int \frac{1}{(4x+5)^2}\mathrm{d}x = -\frac{1}{4} \times \frac{1}{4x+5} + C = -\frac{1}{4(4x+5)} + C.$$

当分母是一个二次函数时, 情况会变得很复杂, 像这样：

$$\int \frac{Ax+B}{ax^2+bx+c}\mathrm{d}x.$$

请注意! 如果分母可以因式分解, 那么就转化为了第一种情况. 下面是前面的一个例子：

$$\int \frac{x+2}{x^2-1}\mathrm{d}x.$$

我们把分母因式分解为 $(x-1)(x+1)$, 这样可得两个被积函数的分母为线性式的积分. 如果是这种情况, 我们就没有必要对分母为二次多项式的被积函数求积分了. 即使上一个例子的分母为 $(4x+5)^2$, 我们也没有必要对二次多项式求积分, 因为它是线性式的平方.

还有什么情况没有考虑到呢？可能的情况是分母的二次多项式不能因式分解. 也就是说, 它的判别式 b^2-4ac 为负. 这类积分的一个例子是

$$\int \frac{x+8}{x^2+6x+13}\mathrm{d}x.$$

它的分母是二次多项式且它的判别式为 $6^2-4\times(13)$, 结果为负. 因为这个分母不能因式分解, 所以对于这道题不能使用刚才的代数运算方法. 我们没有必要去使用任何分部, 所要计算的就是求这个积分. 其做法是：把分母写成平方的形式, 然后再换元. (参见 1.6 节关于配方的讲解.) 在这个例子中, 让我们完成配方：

$$x^2+6x+13=x^2+6x+9+13-9=(x+3)^2+4.$$

所以有

$$\int \frac{x+8}{x^2+6x+13}\mathrm{d}x=\int \frac{x+8}{(x+3)^2+4}\mathrm{d}x.$$

现在换元 $t=x+3$, 所以有 $x=t-3$ 并且 $\mathrm{d}x$=$\mathrm{d}t$：

$$\int \frac{x+8}{x^2+6x+13}\mathrm{d}x=\int \frac{x+8}{(x+3)^2+4}\mathrm{d}x=\int \frac{(t-3)+8}{t^2+4}\mathrm{d}t=\int \frac{t+5}{t^2+4}\mathrm{d}t.$$

第二步是把这个积分分成两个积分, 并把常数 5 移到积分符号的外边, 所以上述积分变为

$$\int \frac{t}{t^2+4}\mathrm{d}t+5\int \frac{1}{t^2+4}\mathrm{d}t.$$

第一个积分就像 18.1 节末尾的那个例子. 分母分子同时乘以 2, 可以发现分母的导数恰恰为分子, 所以我们得到了一个以分母为对数的结果：

$$\int \frac{t}{t^2+4}\mathrm{d}t=\frac{1}{2}\int \frac{2t}{t^2+4}\mathrm{d}t=\frac{1}{2}\ln|t^2+4|+C.$$

实际上, 因为 t^2+4 一直为正, 所以可以把绝对值符号去掉. 现在开始计算第二个积分, 它是

$$5\int \frac{1}{t^2+4}\mathrm{d}t,$$

仅仅需要记住这个有用的公式：

$$\int \frac{1}{t^2+a^2}\mathrm{d}t=\frac{1}{a}\tan^{-1}\left(\frac{t}{a}\right)+C.$$



(你应该通过对等式的右侧求导证明这个结果, 或对等式左边进行换元 t=au.) 无论如何, 当 a =2 时, 这个公式为

$$5\int \frac{1}{t^2+4}\mathrm{d}t = 5\times\frac{1}{2}\tan^{-1}\left(\frac{t}{2}\right)+C.$$

所以, 我们的积分结果为

$$\frac{1}{2}\ln(t^2+4)+\frac{5}{2}\tan^{-1}\left(\frac{t}{2}\right)+C.$$

现在用 x+3 替代 t, 得到最后的结果为

$$\frac{1}{2}\ln((x+3)^2+4)+\frac{5}{2}\tan^{-1}\left(\frac{x+3}{2}\right)+C.$$

表达式 $(x+3)^2+4$ 可以化简为 $x^2+6x+13$, 这正是我们最初的分母. 实际上没有必要展开它, 仅仅回到我们配方的地方, 你就会发现需要的方程了. 所以, 我们最终证明了

$$\int \frac{x+8}{x^2+6x+13}\mathrm{d}x = \frac{1}{2}\ln(x^2+6x+13)+\frac{5}{2}\tan^{-1}\left(\frac{x+3}{2}\right)+C.$$

如果分母的最高项的系数不是 1, 我建议你在配方之前把这个系数提出去. 所以, 计算



$$\int \frac{x+8}{2x^2+12x+26}\mathrm{d}x,$$

把 2 提出来, 把这个积分表达式写为

$$\frac{1}{2}\int \frac{x+8}{x^2+6x+13}\mathrm{d}x.$$

这同我们以前的积分是一样的, 只是在积分符号外边放置了 1/2, 所以它的结果为

$$\frac{1}{4}\ln(x^2+6x+13)+\frac{5}{4}\tan^{-1}\left(\frac{x+3}{2}\right)+C.$$

现在, 我们总结部分分式法, 然后看一个更完整的例子.

18.3.3 方法和一个完整的例子

这是一个关于有理函数积分的完整方法.

第一步 —— **先看分子分母最高项的次数, 如有必要请做除法**. 查看分子的次数是否小于分母次数. 如果是, 那你运气好, 直接进入第二步; 如果不是, 就要做一个多项式的除法了, 然后再进入第二步.

第二步 —— **对分母进行因式分解**. 使用二次公式或猜想一个根, 然后再做除法, 以便因式分解被积函数的分母.

第三步 —— **分部**. 像之前描述的那样, 分别写出带有未知常数的 “分部”. 写下一个像这样的等式：

$$\text{被积函数} = \text{分部}$$

第四步 —— **计算常数的值**. 把方程的两边同时乘以分母, 通过任一方法计算

常数的值：(a) 换掉 x 的值; (b) 系数相等法; 或者结合使用 (a) 和 (b) 两种方法. 现在你能用几个有理函数的和来表示这个被积函数, 这些有理函数可能是分子为常数分母为线性函数的幂, 或者分子为线性函数分母为二次函数.

第五步 —— **求解分母为线性项次幂的积分**. 求解分母是线性函数次幂的积分; 答案将会是对数形式或该线性项的负次幂.

第六步 —— **对分母是二次函数的被积函数求积分**. 对于分母是二次函数且不能因式分解的被积函数求积分, 先配方, 然后换元, 再把它尽可能分解为两个积分. 前者会涉及对数, 而第二个会涉及正切函数的反函数. 如果仅仅有一个积分, 它可能是对数形式又可能是正切函数的反函数形式. 这个公式通常是非常实用的：

$$\boxed{\int \frac{1}{t^2+a^2}\mathrm{d}t = \frac{1}{a}\tan^{-1}\left(\frac{t}{a}\right) + C.}$$

记住你不需要每次都经历完整的六步. 有时可能直接跳到最后一步, 比如上一节的例子：

$$\int \frac{x+8}{x^2+6x+13}\mathrm{d}x$$

在此还有一个很复杂的例子, 它用到了所有的六个步骤：

$$\int \frac{x^5-7x^4+19x^3-10x^2-19x+18}{x^4-5x^3+9x^2}\mathrm{d}x.$$

下面详细介绍怎样应用上述方法解决这个问题.

第一步 —— **先看分子分母最高项的次数, 如有必要请做除法**. 在上面的积分表达式中, 分子最高次数为 5, 分母最高次数为 4. 很讨厌, 我们要做多项式的除法了：

$$\begin{array}{r|l}
 & x-2 \\
x^4-5x^3+9x^2 & x^5-7x^4+19x^3-10x^2-19x+18 \\
 & \underline{x^5-5x^4+9x^3} \\
 & \quad -2x^4+10x^3-10x^2 \\
 & \quad \underline{-2x^4+10x^3-18x^2} \\
 & \qquad\qquad 8x^2-19x+18
\end{array}$$

检查细节! 无论如何, 我们已经看到

$$\frac{x^5-7x^4+19x^3-10x^2-19x+18}{x^4-5x^3+9x^2} = x-2+\frac{8x^2-19x+18}{x^4-5x^3+9x^2}.$$

现在对两边同时求积分可得

$$\int \frac{x^5-7x^4+19x^3-10x^2-19x+18}{x^4-5x^3+9x^2}\mathrm{d}x = \int \left(x-2+\frac{8x^2-19x+18}{x^4-5x^3+9x^2}\right)\mathrm{d}x.$$

等式右侧前两项的积分很容易求, 其积分结果为 $\frac{1}{2}x^2-2x$ (我们将在最后的结果

中 $+C$). 所以现在我们要求解积分

$$\int \frac{8x^2-19x+18}{x^4-5x^3+9x^2}\mathrm{d}x.$$

分子的次数仅仅为 2, 这个数低于分母的次数 4. 我们准备好进入下一步了.

第二步 —— **对分母进行因式分解**. 在分母中我们有一个四次项, 很显然可以把 x^2 提出来. 所以我们把分母因式分解为

$$x^4-5x^3+9x^2=x^2(x^2-5x+9).$$

这个二次项 x^2-5x+9 的判别式为 $(-5)^2-4\times(9)=-11$; 因为它是负的, 所以这个二项式不能因式分解. 我们完成了第二步.

第三步 —— **分部**. 我们已经有两个因子 x^2 和 x^2-5x+9. 不要把第一个因子 x^2 看作二次函数, 而应看作一个线性函数的平方. 为了证明这个观点, 它可以写为 $(x-0)^2$. 所以因子 x^2 产生分部:

$$\frac{A}{x^2}+\frac{B}{x}$$

另外, 因子 x^2-5x+9 产生分部

$$\frac{Cx+D}{x^2-5x+9}.$$

将两者放在一起, 我们有

$$\frac{8x^2-19x+18}{x^2(x^2-5x+9)}=\frac{A}{x^2}+\frac{B}{x}+\frac{Cx+D}{x^2-5x+9}.$$

第四步 —— **计算常数的值**. 现在不得不求常数 A、B、C、D 的值了. 首先把上述等式的两边同时乘以分母 $x^2(x^2-5x+9)$ 得到

$$8x^2-19x+18=A(x^2-5x+9)+Bx(x^2-5x+9)+(Cx+D)x^2.$$

注意, 出现在等式右边每一项的分母部分恰恰没有出现在该项所对应的分部里. 例如, 当用 $x^2(x^2-5x+9)$ 乘以 B/x 时, 约掉了一个 x 得到 $Bx(x^2-5x+9)$.

让我们用一个值替代上述等式的 x. 能消掉这个等式里的大多数项的唯一的 x 值是 $x=0$. 如果把 $x=0$ 代入上述等式, 则

$$18=A(9),$$

所以我们马上就知道 $A=2$. 我们仍然需要求解另外三个常数的值, 所以最好找到这些 x 幂项所对应的系数. 让我们从扩展上述方程开始, 再把 x 的不同次幂合并可得:

$$\begin{aligned}8x^2-19x+18&=Ax^2-5Ax+9A+Bx^3-5Bx^2+9Bx+Cx^3+Dx^2\\&=(B+C)x^3+(A-5B+D)x^2+(-5A+9B)x+9A.\end{aligned}$$

现在可以把 x^3、x^2、x 的系数分别写出了:

$$x^3\text{的系数}: 0 = B + C$$
$$x^2\text{的系数}: 8 = A - 5B + D$$
$$x^1\text{的系数}: -19 = -5A + 9B.$$

注意 x^3 的系数在等式的左侧为 0, 因为等式左边的 $8x^2 - 19x + 18$ 并没有 x^3 项. (顺便说一下, 如果你用系数相等法, 可得到 18=9A, 这同我们把 x =0 代入得到的结果是一样的. 你能说出为什么是这样吗？)

无论怎样, 我们得到一些方程去求解; 从最后一个开始, 然后把 A =2 从后向前代入, 很容易就可得到 $B = -1, D = 1, C = 1$. 把这些值代入第三步的最后一个表达式中有

$$\frac{8x^2 - 19x + 18}{x^2(x^2 - 5x + 9)} = \frac{2}{x^2} + \frac{-1}{x} + \frac{x + 1}{x^2 - 5x + 9}.$$

这就是说

$$\int \frac{8x^2 - 19x + 18}{x^2(x^2 - 5x + 9)}\mathrm{d}x = 2\int \frac{1}{x^2}\mathrm{d}x - \int \frac{1}{x}\mathrm{d}x + \int \frac{x + 1}{x^2 - 5x + 9}\mathrm{d}x.$$

这样我们把一个很复杂的积分化简成了三个简单的积分. 让我们分别对它们求积分.

第五步 —— **求解分母为线性次幂的积分**. 前两个积分是很容易求的：

$$2\int \frac{1}{x^2}\mathrm{d}x - \int \frac{1}{x}\mathrm{d}x = -\frac{2}{x} - \ln|x| + C.$$

所以这道题的第五步真的不麻烦. 但很不走运, 第六步却很烦琐 ……

第六步 —— **对分母是二次函数的被积函数求积分**. 我们需要计算第三个积分, 它是

$$\int \frac{x + 1}{x^2 - 5x + 9}\mathrm{d}x.$$

通过配方可得

$$x^2 - 5x + 9 = \left(x^2 - 5x + \frac{25}{4}\right) + 9 - \frac{25}{4} = \left(x - \frac{5}{2}\right)^2 + \frac{11}{4}. \qquad (**)$$

现在让我们重写积分表达式

$$\int \frac{x + 1}{x^2 - 5x + 9}\mathrm{d}x = \int \frac{x + 1}{(x - \frac{5}{2})^2 + \frac{11}{4}}\mathrm{d}x.$$

我们能用 $t = x - \dfrac{5}{2}$ 做换元. 这时 $x = t + \dfrac{5}{2}$ 且 $\mathrm{d}t = \mathrm{d}x$, 所以这个积分变为

$$\int \frac{t + \frac{5}{2} + 1}{t^2 + \frac{11}{4}}\mathrm{d}t = \int \frac{t + \frac{7}{2}}{t^2 + \frac{11}{4}}\mathrm{d}t,$$

这里以 t 为变量. 现在把它分成两个积分：

$$\int \frac{t}{t^2+\frac{11}{4}}\mathrm{d}t \quad 和 \quad \frac{7}{2}\int \frac{1}{t^2+\frac{11}{4}}\mathrm{d}t.$$

先计算这两个积分中的第一个, 分子分母同时乘以 2 可得

$$\int \frac{t}{t^2+\frac{11}{4}}\mathrm{d}t = \frac{1}{2}\int \frac{2t}{t^2+\frac{11}{4}}\mathrm{d}t = \frac{1}{2}\ln\left|t^2+\frac{11}{4}\right| + C.$$

再一次提醒, 这个绝对值符号是不必要的, 因为 $t^2+\dfrac{11}{4}$ 一定为正. 为把它换回以 x 为变量的状态, 我们需要用 $x-\dfrac{5}{2}$ 替代 t:

$$\frac{1}{2}\ln\left(t^2+\frac{11}{4}\right)+C = \frac{1}{2}\ln\left(\left(x-\frac{5}{2}\right)^2+\frac{11}{4}\right)+C.$$

不要把这个乘法算式展开 —— 仅仅看看上一页我们标记为 (**) 的方程, 当时我们正在配方, 可知这个结果可以化简为 $\dfrac{1}{2}\ln(x^2-5x+9)+C$. 这样, 我们完成了这个积分的第一部分.

我们仍然需要考虑第二部分, 即

$$\frac{7}{2}\int \frac{1}{t^2+\frac{11}{4}}\mathrm{d}t.$$

我们使用公式

$$\int \frac{1}{t^2+a^2}\mathrm{d}t = \frac{1}{a}\tan^{-1}\left(\frac{t}{a}\right)+C,$$

其中 $a=\sqrt{11/4}$, 事实上它等于 $\sqrt{11}/2$:

$$\frac{7}{2}\int \frac{1}{t^2+\frac{11}{4}}\mathrm{d}t = \frac{7}{2}\times\frac{2}{\sqrt{11}}\tan^{-1}\left(\frac{t}{\sqrt{11}/2}\right)+C = \frac{7}{\sqrt{11}}\tan^{-1}\left(\frac{2t}{\sqrt{11}}\right)+C.$$

现在把 $t=x-\dfrac{5}{2}$ 再次代入可得

$$\frac{7}{\sqrt{11}}\tan^{-1}\left(\frac{2x-5}{\sqrt{11}}\right)+C.$$

最后两个积分给出了第六步的最终答案:

$$\int \frac{x+1}{x^2-5x+9}\mathrm{d}x = \frac{1}{2}\ln(x^2-5x+9)+\frac{7}{\sqrt{11}}\tan^{-1}\left(\frac{2x-5}{\sqrt{11}}\right)+C.$$

猜想我们将要做什么? 我们正准备把得到的所有结果放到一起! 前 4 步中我们得到

$$\int \frac{x^5-7x^4+19x^3-10x^2-19x+18}{x^4-5x^3+9x^2}\mathrm{d}x$$
$$= \int\left(x-2+\frac{2}{x^2}-\frac{1}{x}+\frac{x+1}{x^2-5x+9}\right)\mathrm{d}x.$$

这是完整的部分分式分解形式. 现在再用第五步和第六步计算这个积分, 上述的积分结果为

$$\frac{x^2}{2} - 2x - \frac{2}{x} - \ln|x| + \frac{1}{2}\ln(x^2 - 5x + 9) + \frac{7}{\sqrt{11}}\tan^{-1}\left(\frac{2x-5}{\sqrt{11}}\right) + C.$$



我们终于解决了这个复杂的例子. 它确实是很复杂的, 但是如果你能解答这么难的问题, 那么对于简单的问题你就游刃有余了. 作为一个练习, 看你明天不看这几页是否能独立做出这道题目.

第 19 章　积分的方法 II

在这一章中, 我们将继续介绍积分方法 —— 求解涉及三角函数的积分方法. 有时, 我们需要使用三角恒等式解决一些问题; 有时题目中没有三角函数, 而在计算过程中需要使用三角换元法. 介绍完三角函数的积分方法之后, 我们将要总结一下本章和上一章讲述的积分方法. 所以, 在这一章中我们将要学习如下知识点:

- 关于三角恒等式的积分;
- 关于三角函数的幂以及约化公式的积分;
- 关于三角换元法的积分;
- 关于所学习过的所有积分方法的总结.

19.1　应用三角恒等式的积分

有三大类型的三角恒等式, 它们在积分计算中非常有用. 第一大类型是关于 $\cos(2x)$ 的倍角公式. 在 2.4 节中, 我们知道 $\cos(2x) = 2\cos^2(x) - 1$, 也知道 $\cos(2x) = 1 - 2\sin^2(x)$. (请记住, 其中的一个可以由另一个应用公式 $\sin^2(x) + \cos^2(x) = 1$ 推导出来.) 该公式在积分计算中最好的应用地方是当被积函数中出现 $\sin^2(x)$ 和 $\cos^2(x)$ 时. 所以, 我们有

$$\boxed{\cos^2(x) = \frac{1}{2}(1 + \cos(2x))} \quad 和 \quad \boxed{\sin^2(x) = \frac{1}{2}(1 - \cos(2x))}$$

这两个公式很值得记住! 具体而言, 如果需要求 $1 + \cos$ (任何值) 或 $1 - \cos$(任何值) 的平方根, 那这两个公式就派上用场了. 例如,



$$\int_0^{\pi/2} \sqrt{1 - \cos(2x)}\mathrm{d}x$$

这道题看起来很麻烦, 但实际上

$$\int_0^{\pi/2} \sqrt{1 - \cos(2x)}\mathrm{d}x = \int_0^{\pi/2} \sqrt{2\sin^2(x)}\mathrm{d}x$$

可以用刚才的第二个方框里的公式导出. (我们在使用这个公式前需要乘以 2.) 然而, 如果直接用 $\sqrt{2}\sin(x)$ 替代 $\sqrt{2\sin^2(x)}$ 是很鲁莽的, 我们需要做个检测. A 的平方根不一定是 A 本身, 而是 $|A|$. 所以上述积分变为

$$\sqrt{2}\int_0^{\pi/2} |\sin(x)|\mathrm{d}x.$$

幸运的是, 当 x 在 0 到 $\pi/2$ 之间时, $\sin(x)$ 的值一直大于或等于零, 所以我们最终可把绝对值符号去掉! 我们已经导出

$$\sqrt{2}\int_0^{\pi/2}\sin(x)\mathrm{d}x,$$

剩余的计算留给你去做, 它的结果为 $\sqrt{2}$.

有时你需要更灵活. 考虑

$$\int_\pi^{2\pi}\sqrt{1+\cos(x)}\mathrm{d}x.$$

看起来我们需要使用刚才第一个方框里的公式, 但原始公式是 $1+\cos(2x)$, 而我们的被积函数是 $1+\cos(x)$. 没问题. 如果你用 $x/2$ 去替代 x, 然后再乘以 2, 就得到

$$2\cos^2\left(\frac{x}{2}\right)=1+\cos(x).$$

这正是我们想要的! 检验这个:

$$\int_\pi^{2\pi}\sqrt{1+\cos(x)}\mathrm{d}x=\int_\pi^{2\pi}\sqrt{2\cos^2\left(\frac{x}{2}\right)}\mathrm{d}x=\sqrt{2}\int_\pi^{2\pi}\left|\cos\left(\frac{x}{2}\right)\right|\mathrm{d}x.$$

现在我们得非常小心! 当 x 在 π 到 2π 之间时, $x/2$ 是在 $\pi/2$ 到 π 之间, 但 $\cos(x)$ 在区间 $[\pi/2,\pi]$ 上是小于等于零的 (可以画图像校验). 所以上述积分实际上等于

$$\sqrt{2}\int_\pi^{2\pi}\left(-\cos\left(\frac{x}{2}\right)\right)\mathrm{d}x;$$

剩余的计算工作留给你去做, 它的结果为 $2\sqrt{2}$. 顺便说一下, 如果你错误地用 $\cos(x/2)$ 而不是 $-\cos(x/2)$ 替代 $|\cos(x/2)|$, 那么得到的答案将会是 $-2\sqrt{2}$. 这是不正确的, 因为原始的被积函数 $\sqrt{1+\cos(x)}$ 一直为正的, 所以这个积分的结果也应该为正.

让我们接下来讨论第二大类型的三角恒等式, 它们是毕达哥拉斯恒等式:

$$\boxed{\sin^2(x)+\cos^2(x)=1}\quad\boxed{\tan^2(x)+1=\sec^2(x)}\quad\boxed{1+\cot^2(x)=\csc^2(x)}$$

如 2.4 节所述, 这些等式对于所有的 x 都适用. 有时它们是很有帮助的. 例如,

$$\int_0^\pi\sqrt{1-\cos^2(x)}\mathrm{d}x$$

应该被写为

$$\int_0^\pi\sqrt{\sin^2(x)}\mathrm{d}x=\int_0^\pi|\sin(x)|\mathrm{d}x.$$

因为当 x 在 0 到 π 之间时 $\sin(x)\geqslant 0$, 我们可以去掉绝对值符号, 写为

$$\int_0^\pi\sin(x)\mathrm{d}x,$$

结果为 2. (你自己计算一下!) 把这个例子 $\int_0^\pi\sqrt{1-\cos^2(x)}\mathrm{d}x$ 和我们刚才做过的例

子 $\int_0^\pi \sqrt{1-\cos(x)}\mathrm{d}x$ 进行比较. 它们看起来可能很相似, 但所使用的三角恒等式是不同的.

有时你不得不应用一些技巧才能使用上述公式. 如果你在一个积分的分母中看到 1+trig(x) 或 1−trig(x), 其中 trig 是三角函数的意思 (可能是正弦、余弦、正割或余割), 那么可以考虑用这个积分表达式乘以与其分母共轭的表达式. 例如, 计算





$$\int \frac{1}{\sec(x)-1}\mathrm{d}x,$$

分子分母同时乘以分母的共轭表达式, 在此是 $\sec(x)+1$. 也就是

$$\int \frac{1}{\sec(x)-1}\mathrm{d}x = \int \frac{1}{\sec(x)-1} \times \frac{\sec(x)+1}{\sec(x)+1}\mathrm{d}x.$$

现在可以在积分的分母表达式中使用平方差公式 $(a-b)(a+b)=a^2-b^2$, 把这个积分写为

$$\int \frac{\sec(x)+1}{\sec^2(x)-1}\mathrm{d}x.$$

根据刚才方框里的公式, 分母恰恰就是 $\tan^2(x)$. 使用这个结果重写这个积分, 然后再把它分成两个积分, 我们发现原始积分变为

$$\int \frac{\sec(x)+1}{\tan^2(x)}\mathrm{d}x = \int \frac{\sec(x)}{\tan^2(x)}\mathrm{d}x + \int \frac{1}{\tan^2(x)}\mathrm{d}x.$$

第一个积分看起来不容易计算, 但可以用正弦和余弦的形式来表示它. 具体情况是

$$\int \frac{\sec(x)}{\tan^2(x)}\mathrm{d}x = \int \frac{1/\cos(x)}{\sin^2(x)/\cos^2(x)}\mathrm{d}x = \int \frac{\cos(x)}{\sin^2(x)}\mathrm{d}x.$$



下一步是换元 $t=\sin(x)$, 因为在分子中 $\mathrm{d}t=\cos(x)\mathrm{d}x$. 试试看你能得到什么. 更简单的方式是把 $\cos(x)/\sin^2(x)$ 改写成 $\csc(x)\cot(x)$, 所以

$$\int \frac{\cos(x)}{\sin^2(x)}\mathrm{d}x = \int \csc(x)\cot(x)\mathrm{d}x = -\csc(x)+C,$$

因为 $\csc(x)$ 的导数为 $-\csc(x)\cot(x)$. 现在我们需要计算第二个积分：

$$\int \frac{1}{\tan^2(x)}\mathrm{d}x.$$

没问题, 把这个积分表达式改写为 $\int \cot^2(x)\mathrm{d}x$, 这时使用刚才方框里的另一个三角恒等式, 把这个表达式改为

$$\int (\csc^2(x)-1)\mathrm{d}x = -\cot(x)-x+C.$$

(你还能记住 $\csc^2(x)$ 的积分结果吗? 它与 $\sec^2(x)$ 的积分很相似, $\sec^2(x)$ 的积分的结果为 $\tan(x)+C$. 仅仅在转换过程中加一个“co-” (“余” 字) 和一个负号, 就得到了 $\csc^2(x)$ 形式的积分结果!) 我们把这两个积分结果放到一起得到

$$\int \frac{1}{\sec(x)-1}\mathrm{d}x = -\csc(x) - \cot(x) - x + C.$$

以上计算确实需要一些技巧.

让我们看看第三大类型的三角恒等式, 它叫作积化和差公式：

$$\boxed{\begin{aligned} \cos(A)\cos(B) &= \frac{1}{2}(\cos(A-B)+\cos(A+B)) \\ \sin(A)\sin(B) &= \frac{1}{2}(\cos(A-B)-\cos(A+B)) \\ \sin(A)\cos(B) &= \frac{1}{2}(\sin(A-B)+\sin(A+B)) \end{aligned}}$$

这些公式确实不容易记住. 实际上它们都从表达式 $\cos(A \pm B)$ 和 $\sin(A \pm B)$ 而来 (可在 2.4 节找到相关知识), 如果你已经掌握了这些公式, 就可以很容易地把它们转换过来. 这些公式对于

$$\int \cos(3x)\sin(19x)\mathrm{d}x$$

这类积分是必要的. 实际上, 我们可以使用上面的第三个公式解决这个问题, $A = 19x$ 和 $B = 3x$. (千万不要让 cos 和 sin 的顺序愚弄了你! 该积分同 $\int \sin(19x)\cos(3x)\mathrm{d}x$ 是一样的.) 使用这个公式可得

$$\begin{aligned} \int \cos(3x)\sin(19x)\mathrm{d}x &= \frac{1}{2}\int (\sin(19x-3x)+\sin(19x+3x))\mathrm{d}x \\ &= \frac{1}{2}\int (\sin(16x)+\sin(22x))\mathrm{d}x \\ &= \frac{1}{2}\left(-\frac{\cos(16x)}{16} - \frac{\cos(22x)}{22}\right) + C \\ &= -\frac{\cos(16x)}{32} - \frac{\cos(22x)}{44} + C. \end{aligned}$$

19.2 关于三角函数的幂的积分

现在我们将要研究怎样求解被积函数是三角函数的幂的形式的积分, 例如求解 $\int \cos^7(x)\sin^{10}(x)\mathrm{d}x$ 或 $\int \sec^6(x)\mathrm{d}x$. 遗憾的是, 被积函数中的三角函数的类型不同, 求解积分所要求的积分技巧也不同. 所以我们来分别讨论它们.

19.2.1 sin 或 cos 的幂

刚才的例子 $\int \cos^7(x)\sin^{10}(x)\mathrm{d}x$ 就属于这种类型. 这里有一个黄金法则：如果 $\sin(x)$ 或 $\cos(x)$ 其中一个的幂是奇数, 那就一定要抓住它 —— 它是你的朋友! (如果两个都为奇数, 把幂低的那个选做你的朋友.) 如果你已经抓住了奇次幂, 这时需

要做的是拿出一项同 $\mathrm{d}x$ 放在一起, 再用下列公式中的一个处理剩下的项 (现在是偶次幂了)

$$\boxed{\cos^2(x) = 1 - \sin^2(x)} \quad 或 \quad \boxed{\sin^2(x) = 1 - \cos^2(x)}$$

注意这两个公式就是 $\sin^2(x) + \cos^2(x) = 1$ 的另一种写法. 我们来看怎样使用这个方法. 在 $\int \cos^7(x)\sin^{10}(x)\mathrm{d}x$ 中, 注意 7 是奇数, 于是我们得到了 $\cos^7(x)$, 这时需要移出一个 $\cos(x)$ 并把它和 $\mathrm{d}x$ 放到一起. 我们得到

$$\int \cos^7(x)\sin^{10}(x)\mathrm{d}x = \int \cos^6(x)\sin^{10}(x)\cos(x)\mathrm{d}x.$$

这又能怎样呢? 很好, 我们需要处理剩下的 $\cos^6(x)$. 现在 6 是偶数, 所以可以写为 $\cos^6(x) = \left(\cos^2(x)\right)^3 = \left(1 - \sin^2(x)\right)^3$, 这样该积分为

$$\int (1 - \sin^2(x))^3 \sin^{10}(x)\cos(x)\mathrm{d}x.$$

现在如果设 $t = \sin(x)$, 则 $\mathrm{d}t = \cos(x)\mathrm{d}x$, 所以很容易用 t 为变量来表示这个积分:

$$\int (1 - t^2)^3 t^{10}\mathrm{d}t = \int (1 - 3t^2 + 3t^4 - t^6)t^{10}\mathrm{d}t = \int (t^{10} - 3t^{12} + 3t^{14} - t^{16})\mathrm{d}t,$$

结果为

$$\frac{t^{11}}{11} - \frac{3t^{13}}{13} + \frac{t^{15}}{5} - \frac{t^{17}}{17} + C.$$

再把它换回以 x 为变量的积分, 就得到了答案:

$$\int \cos^7(x)\sin^{10}(x)\mathrm{d}x = \frac{\sin^{11}(x)}{11} - \frac{3\sin^{13}(x)}{13} + \frac{\sin^{15}(x)}{5} - \frac{\sin^{17}(x)}{17} + C.$$

你看到了吧, 借用一个 $\cos(x)$ 来帮助我们改变被积函数把 $\cos(x)$ 和 $\mathrm{d}x$ 结合在一起做换元 $t = \sin(x)$, 从而使被积函数仅仅关于 $\sin(x)$:

如果它们的幂都不是奇数该怎么办呢? 很好, 如果它们的幂都为偶数, 例如 $\int \cos^2(x)\sin^4(x)\mathrm{d}x$, 那应该使用倍角公式. 我们在上一节中见过这些公式, 这里将再一次用到它们了:

$$\boxed{\cos^2(x) = \frac{1}{2}(1 + \cos(2x))} \quad 和 \quad \boxed{\sin^2(x) = \frac{1}{2}(1 - \cos(2x)).}$$

直接用这两个公式做替代, 可以看到关于 cos 的幂的更简单的被积函数. 这时你可以使用刚才的计算方法, 看积分的每一部分是奇次还是偶次. 在这个的例子中, 我们需要把 $\sin^4(x)$ 用 $\left(\sin^2(x)\right)^2$ 来表示, 所以有

$$\int \cos^2(x)\sin^4(x)\mathrm{d}x = \int \frac{1}{2}(1 + \cos(2x))\left(\frac{1}{2}(1 - \cos(2x))\right)^2 \mathrm{d}x.$$

现在把它展开得到

$$\frac{1}{8}\int(1-\cos(2x)-\cos^2(2x)+\cos^3(2x))\mathrm{d}x.$$

我们需要把这个积分分解为四个单独的积分. 我们暂时先不要考虑积分符号前面的 1/8 或负号. 前面两个积分很容易计算, 因为 $\int 1\mathrm{d}x = x + C$, $\int\cos(2x)\mathrm{d}x = \frac{1}{2}\sin(2x)+C$. 我们怎样才能计算 $\int\cos^2(2x)\mathrm{d}x$ 呢? 这是一个偶次方, 我们需要再次使用倍角公式, 但需要用 $2x$ 替代 x:

$$\int\cos^2(2x)\mathrm{d}x = \int\frac{1}{2}(1+\cos(4x))\mathrm{d}x = \frac{1}{2}\left(x+\frac{1}{4}\sin(4x)\right)+C.$$

那么 $\int\cos^3(2x)\mathrm{d}x$ 又该怎样计算呢? 很好, 现在它是奇次的 (也就是 3), 所以我们已经知道怎么做了! 把这个积分写为 $\int\cos^2(2x)\cos(2x)\mathrm{d}x$, 然后用 $(1-\sin^2(2x))$ 替代 $\cos^2(2x)$. 用 $t=\sin(2x)$ 换元, 这样就有 $\mathrm{d}t = 2\cos(2x)\mathrm{d}x$, 所以 $\int\cos^3(2x)\mathrm{d}x$ 这个积分为

$$\begin{aligned}\int(1-\sin^2(2x))\cos(2x)\mathrm{d}x &= \frac{1}{2}\int(1-t^2)\mathrm{d}t = \frac{1}{2}\left(t-\frac{t^3}{3}\right)+C\\ &= \frac{\sin(2x)}{2}-\frac{\sin^3(2x)}{6}+C.\end{aligned}$$

(休息一下.) 现在把这些综合到一起再化简, 你会发现我们得到了

$$\begin{aligned}&\int\cos^2(x)\sin^4(x)\mathrm{d}x\\ &= \frac{1}{8}\left(x-\frac{\sin(2x)}{2}-\frac{x}{2}-\frac{\sin(4x)}{8}+\frac{\sin(2x)}{2}-\frac{\sin^3(2x)}{6}\right)+C\\ &= \frac{x}{16}-\frac{\sin(4x)}{64}-\frac{\sin^3(2x)}{48}+C.\end{aligned}$$

要确保你自己也能做出来.

19.2.2　tan 的幂

考虑 $\int\tan^n(x)\mathrm{d}x$, 其中 n 是整数. 我们先研究前几种情况. 当 $n=1$ 时, 我们需要知道怎样计算 $\int\tan x\mathrm{d}x$. 这是一个标准的积分, 可以通过设 $t=\cos(x)$ 来解答, 注意 $\mathrm{d}t=-\sin(x)\mathrm{d}x$:

$$\int\tan(x)\mathrm{d}x = \int\frac{\sin(x)}{\cos(x)}\mathrm{d}x = -\int\frac{\mathrm{d}t}{t} = -\ln(t)+C = -\ln|\cos(x)|+C.$$

这个答案也可以被写为 $\ln|\sec(x)|+C$. (为什么呢?)

当 $n=2$ 时情况又是怎样呢? 对于这种情况, 我们有必要使用毕达哥拉斯恒等式:

$$\boxed{\tan^2(x)=\sec^2(x)-1}$$

我们在上一节见过这个公式. 所以有

$$\int \tan^2(x)\mathrm{d}x = \int (\sec^2(x) - 1)\mathrm{d}x = \tan(x) - x + C.$$

对于更高次幂 ($n \geqslant 3$), 就不得不把 $\tan^2(x)$ 先提出来再改写为 $(\sec^2(x) - 1)$. 这样就有了两个积分. 前面的积分可以通过设 $t = \tan(x)$ 来计算并使用 $\mathrm{d}t = \sec^2(x)\mathrm{d}x$. 第二个积分是 $\tan(x)$ 的更低次幂, 所要做的是重复这个方法. 例如, 怎样计算 $\int \tan^6(x)\mathrm{d}x$? 让我们看看:

$$\begin{aligned}\int \tan^6(x)\mathrm{d}x &= \int \tan^4(x)\tan^2(x)\mathrm{d}x = \int \tan^4(x)(\sec^2(x) - 1)\mathrm{d}x \\ &= \int \tan^4(x)\sec^2(x)\mathrm{d}x - \int \tan^4(x)\mathrm{d}x.\end{aligned}$$

现在我们需要计算这两个积分. 为计算第一个积分, 我们设 $t = \tan(x)$; 像我们说过的那样 $\mathrm{d}t = \sec^2(x)\mathrm{d}x$. 这样给出了

$$\int \tan^4(x)\sec^2(x)\mathrm{d}x = \int t^4\mathrm{d}t = \frac{t^5}{5} + C = \frac{\tan^5(x)}{5} + C.$$

现在, 第二个积分为 $\int \tan^4(x)\mathrm{d}x$, 所以还得重复这个过程. 提出一个 $\tan^2(x)$ 因子, 然后把它改为 $(\sec^2(x) - 1)$:

$$\begin{aligned}\int \tan^4(x)\mathrm{d}x &= \int \tan^2(x)\tan^2(x)\mathrm{d}x = \int \tan^2(x)(\sec^2(x) - 1)\mathrm{d}x \\ &= \int \tan^2(x)\sec^2(x)\mathrm{d}x - \int \tan^2(x)\mathrm{d}x.\end{aligned}$$

再一次地, 我们有了两个积分. 为计算第一个, 设 $t = \tan(x)$, 所以有 $\mathrm{d}t = \sec^2(x)\mathrm{d}x$. (熟悉吗?) 所以

$$\int \tan^2(x)\sec^2(x)\mathrm{d}x = \int t^2\mathrm{d}t = \frac{t^3}{3} + C = \frac{\tan^3(x)}{3} + C.$$

与此同时, 我们看到

$$\int \tan^2(x)\mathrm{d}x = \int (\sec^2(x) - 1)\mathrm{d}x = \tan(x) - x + C.$$

把这些计算结果合并到一起 (记住不要忘记负号), 得到

$$\int \tan^6(x)\mathrm{d}x = \frac{\tan^5(x)}{5} - \frac{\tan^3(x)}{3} + \tan(x) - x + C.$$

确实有些复杂. 但是, 还有更复杂呢.

19.2.3 sec 的幂

这种类型的积分确实很难算, 只有当 $n = 2$ (即 $\int \sec^2(x)\mathrm{d}x$) 时容易计算. 我们从一次幂 $\int \sec(x)\mathrm{d}x$ 开始. 计算这个积分有许多种方法. 最容易的方法要用到一个很巧妙的技巧. 这个技巧很节省时间, 值得一记. 不走运的是, 这种技巧完全超过了正

常人的思维, 难以在第一时间想到. 这个方法是分子分母同时乘以 $(\sec(x)+\tan(x))$. 看看这个计算过程, 它真的很奇妙:

$$\begin{aligned}\int \sec(x)\mathrm{d}x &= \int \sec(x)\times\frac{\sec(x)+\tan(x)}{\sec(x)+\tan(x)}\mathrm{d}x = \int \frac{\sec^2(x)+\sec(x)\tan(x)}{\sec(x)+\tan(x)}\mathrm{d}x\\ &= \ln|\sec(x)+\tan(x)|+C,\end{aligned}$$

因为分母 $(\sec(x)+\tan(x))$ 的导数恰恰等同于分子.

$\sec(x)$ 的二次幂该怎样计算呢? 这个不需要太费力气:

$$\int \sec^2(x)\mathrm{d}x = \tan(x)+C.$$

很容易计算. 不幸的是, 更高次幂就很难计算了. 不过, 基本思想是把 $\sec^2(x)$ 提出来 (这同我们前面处理 $\tan(x)$ 的幂很相似), 用分部积分法, 应用 $\mathrm{d}v=\sec^2(x)\mathrm{d}x$ 并把 u 设为余下的 $\sec(x)$ 次幂. 也就是说, $v=\tan(x)$ (记住在这里不需要常数项). 当用分部积分法时, 自然会得到一个新积分; 被积函数应该是一个 $\sec(x)$ 的更低次幂乘以 $\tan^2(x)$. 我们需要再一次使用 $\tan^2(x)=\sec^2(x)-1$ 并得到两个积分, 而其中一个就是原始积分的倍数! 你需要把这个放回等式的左边. 另一个是关于 $\sec(x)$ 的更低次幂, 你需要重复整个过程直到剩下 $\int\sec(x)\mathrm{d}x$ 或 $\int\sec^2(x)\mathrm{d}x$, 这两个积分的结果我们已经知道了.

这是一个技术解释, 我们来看一个很难对付的例子: 计算 $\int\sec^6(x)\mathrm{d}x$. 我们先把 $\sec^2(x)$ 提出来, 即

$$\int \sec^6(x)\mathrm{d}x = \int \sec^4(x)\sec^2(x)\mathrm{d}x.$$

现在, 使用分部积分法, 设 $u=\sec^4(x), \mathrm{d}v=\sec^2(x)\mathrm{d}x$. 通过对 u 求导和对 $\mathrm{d}v$ 求积分, 我们得到

$$\mathrm{d}u = 4\sec^3(x)\sec(x)\tan(x)\mathrm{d}x = 4\sec^4(x)\tan(x)\mathrm{d}x \text{ 和 } v=\tan(x).$$

现在通过分部积分法得到

$$\int u\mathrm{d}v = uv - \int v\mathrm{d}u$$

$$\int \sec^4(x)\overbrace{\sec^2(x)\mathrm{d}x}^{\mathrm{d}v} = \sec^4(x)\tan(x) - \int \tan(x)\overbrace{4\sec^4(x)\tan(x)\mathrm{d}x}^{\mathrm{d}u}.$$

来看看等式右侧的积分, 它可以写为

$$\begin{aligned}4\int \sec^4(x)\tan^2(x)\mathrm{d}x &= 4\int \sec^4(x)(\sec^2(x)-1)\mathrm{d}x\\ &= 4\left(\int \sec^6(x)\mathrm{d}x - \int \sec^4(x)\mathrm{d}x\right).\end{aligned}$$

把这些放到一起, 我们有

$$\int \sec^6(x)\mathrm{d}x = \sec^4(x)\tan(x) - 4\int \sec^6(x)\mathrm{d}x + 4\int \sec^4(x)\mathrm{d}x.$$

下一步令人振奋：把等式右侧的第一个积分移到等式左侧：

$$5\int \sec^6(x)\mathrm{d}x = \sec^4(x)\tan(x) + 4\int \sec^4(x)\mathrm{d}x.$$

等式两侧同时除以 5 可得

$$\int \sec^6(x)\mathrm{d}x = \frac{1}{5}\sec^4(x)\tan(x) + \frac{4}{5}\int \sec^4(x)\mathrm{d}x.$$



做完了吗? 还没有, 我们仍然需要计算 $\int \sec^4(x)\mathrm{d}x$. 我们不得不重复刚才的全过程. 这正是你需要重复上述步骤的地方. 如果你没有计算错误, 会得到

$$\int \sec^4(x)\mathrm{d}x = \frac{1}{3}\sec^2(x)\tan(x) + \frac{2}{3}\int \sec^2(x)\mathrm{d}x.$$

现在我们需要计算 $\int \sec^2(x)\mathrm{d}x$, 这达到了我们力所能及的程度 —— 它的结果是 $\tan(x) + C$, 我们以前见过的. 再把这些都合并到一起, 得到

$$\begin{aligned}\int \sec^6(x)\mathrm{d}x &= \frac{1}{5}\sec^4(x)\tan(x) + \frac{4}{5}\left(\frac{1}{3}\sec^2(x)\tan(x) + \frac{2}{3}\tan(x)\right) + C\\ &= \frac{1}{5}\sec^4(x)\tan(x) + \frac{4}{15}\sec^2(x)\tan(x) + \frac{8}{15}\tan(x) + C.\end{aligned}$$

很好, 尽管我们费了很大的劲, 但已经算出了结果. 看, 解决带有 $\tan(x)$ 和 $\sec(x)$ 的幂的习题的基本思想是：先降 2 次幂, 然后重复计算; 继续计算, 直到降为一次幂或二次幂, 这样我们就可以直接计算了. 顺便想一想, 怎样计算



$$\int \frac{\mathrm{d}x}{\cos^6(x)}?$$

当然, 我们可以把它写为 $\int \sec^6(x)\mathrm{d}x$(我们刚刚已经计算出结果了!). 那么又怎样计算



$$\int \frac{\sin^2(x)}{\cos^3(x)}\mathrm{d}x?$$

分子可以改写为 $1 - \cos^2(x)$, 然后分成两个积分：

$$\int \frac{\sin^2(x)}{\cos^3(x)}\mathrm{d}x = \int \frac{1-\cos^2(x)}{\cos^3(x)}\mathrm{d}x = \int \sec^3(x)\mathrm{d}x - \int \sec(x)\mathrm{d}x.$$

现在使用上述方法, 我们就可以求出关于 $\sec(x)$ 的次幂的积分了.

19.2.4 cot 的幂

我们可以用解决 $\tan(x)$ 的幂的方法来解决这类问题. 可以使用毕达哥拉斯恒等式把 $\cot^2(x)$ 改写：

$$\boxed{\cot^2(x) = \csc^2(x) - 1}$$



当设 $t=\cot(x)$ 时, 有 $\mathrm{d}t=-\csc^2(x)\mathrm{d}x$. 请注意, 不要忘记负号! 现在多做一些题目来练习, 例如计算 $\int\cot^6(x)\mathrm{d}x$. 将这个结果和 19.2.2 节中 $\int\tan^6(x)\mathrm{d}x$ 的结果进行比较. 你会发现它们是非常相似的.

19.2.5 csc 的幂



计算这个就和计算 $\sec(x)$ 的幂一样. 可以把 $\csc^2(x)$ 提出来, 然后用分部积分法, 应用 $\mathrm{d}v=\csc^2(x)\mathrm{d}x$. 请注意: $v=-\cot(x)$, 而 $\mathrm{d}u$ 也有一个负号, 这是你需要注意的地方. 再一次提醒要多做些练习. 例如计算 $\int\csc^6(x)\mathrm{d}x$, 将这个结果和 $\int\sec^6(x)\mathrm{d}x$ 的结果相比较, 你会看到更多相似的地方.

19.2.6 约化公式



前四节的方法都是把三角函数的幂降低 2 次, 然后重复计算. 例如 在 19.2.2 节中, 通过提出 $\tan^2(x)$ 然后用 $\sec^2(x)-1$ 替代它来求解 $\tan(x)$ 的幂的积分. 让我们试着总结一下这个方法. 首先来计算 $\int\tan^n(x)\mathrm{d}x$, 我们给它起个名字: I_n(对于整数 n). 也就是说,

$$I_n=\int\tan^n(x)\mathrm{d}x.$$

我们已经知道

$$I_0=\int\tan^0(x)\mathrm{d}x=\int 1\mathrm{d}x=x+C,$$

$$I_1=\int\tan(x)\mathrm{d}x=-\ln|\cos(x)|+C.$$

当 $n\geqslant 2$ 时, 我们可以从 $\tan^n(x)$ 中提取 $\tan^2(x)$, 这样就剩下了 $\tan^{n-2}(x)$; 这时可以使用三角函数恒等式把这个积分分开:

$$\begin{aligned}I_n&=\int\tan^n(x)\mathrm{d}x=\int\tan^{n-2}(x)\tan^2(x)\mathrm{d}x=\int\tan^{n-2}(x)(\sec^2(x)-1)\mathrm{d}x\\&=\int\tan^{n-2}(x)\sec^2(x)\mathrm{d}x-\int\tan^{n-2}(x)\mathrm{d}x.\end{aligned}$$

等式右侧的第二个积分 $\int\tan^{n-2}(x)\mathrm{d}x$ 就是 I_{n-2}; 对于第一个, 如果设 $t=\tan(x)$, 会得到 $\mathrm{d}t=\sec^2(x)\mathrm{d}x$, 这个积分就变为 $\int t^{n-2}\mathrm{d}t$, 它的结果为 $t^{n-1}/(n-1)+C$. 用 $\tan(x)$ 回代 t, 这样我们证明了

$$I_n=\frac{1}{n-1}\tan^{n-1}(x)-I_{n-2}.$$

我们没有必要写常数, 因为 I_n 和 I_{n-2} 都是不定积分. 上述方程叫作约化公式, 因为它把整数 n 降到一个更小的数 $n-2$.

让我们看看怎样使用这个公式计算 $\int\tan^6(x)\mathrm{d}x$, 即 I_6. 把 $n=6$ 代入约化公式, 有

$$I_6 = \frac{1}{5}\tan^5(x) - I_4.$$

很好, 我们需要知道 I_4. 让我们再次使用约化公式, 这次 n =4：

$$I_4 = \frac{1}{3}\tan^3(x) - I_2.$$

再用一次约化公式, n =2：

$$I_2 = \frac{1}{1}\tan^1(x) - I_0 = \tan(x) - x + C,$$

在这个结果里, 我们使用了 I_0. 现在我们知道了 I_2, 可以回去求解 I_4 了：

$$I_4 = \frac{1}{3}\tan^3(x) - I_2 = \frac{1}{3}\tan^3(x) - \tan(x) + x + C.$$

最后可以求解要计算的积分 I_6 了：

$$\int \tan^6(x)\mathrm{d}x = I_6 = \frac{1}{5}\tan^5(x) - I_4 = \frac{1}{5}\tan^5(x) - \frac{1}{3}\tan^3(x) + \tan(x) - x + C.$$

这同 19.2.2 节中的答案一样. 现在把这个方法应用到求解正割、余割和余切的幂的积分当中, 只需要把它们重写为约化公式.

这个方法对于定积分也适用. 例如计算定积分 $\int_0^{\pi/2}\cos^8(x)\mathrm{d}x$ 的值. 如 19.2.1 节所述, 你应该使用倍角公式, 但应用于这道题可能会很麻烦. (不信你可以试试!) 相反, 我们设



$$I_n = \int_0^{\pi/2}\cos^n(x)\mathrm{d}x,$$

要记住我们最后要求 I_8. 现在的技巧是我们需要提出一个因子 $\cos(x)$, 像这样

$$I_n = \int_0^{\pi/2}\cos^n(x)\mathrm{d}x = \int_0^{\pi/2}\cos^{n-1}(x)\cos(x)\mathrm{d}x.$$

现在使用分部积分法, 设 $u = \cos^{n-1}(x), \mathrm{d}v = \cos(x)\mathrm{d}x$. 这就是说, $v = \sin(x)$. (更多关于分部积分法的内容请参考 18.2 节.) 请你证明



$$I_n = \cos^{n-1}(x)\sin(x)\Big|_0^{\pi/2} + \int_0^{\pi/2}(n-1)\cos^{n-2}(x)\sin^2(x)\mathrm{d}x.$$

如果 $n \geqslant 2$, 这时在等式右侧第一项的结果是 0, 因为 $\cos(\pi/2) = 0$, $\sin(0) = 0$. 另一方面, 在积分中, 我们可以用 $1 - \cos^2(x)$ 替代 $\sin^2(x)$, 得到

$$\begin{aligned} I_n &= \int_0^{\pi/2}(n-1)\cos^{n-2}(x)(1-\cos^2(x))\mathrm{d}x \\ &= (n-1)\int_0^{\pi/2}\cos^{n-2}(x)\mathrm{d}x - (n-1)\int_0^{\pi/2}\cos^n(x)\mathrm{d}x. \end{aligned}$$

我们得到了什么? 很好, 请注意后两个积分分别是 I_{n-2} 和 I_n. 所以

$$I_n = (n-1)I_{n-2} - (n-1)I_n.$$

通过把等式两端同时加 $(n-1)I_n$ 再除以 n, 我们得到了这个约化公式：

$$I_n = \frac{n-1}{n} I_{n-2}.$$

这应该使我们的计算更容易! 我们要求 I_8 的解, 所以会一次又一次地使用上述公式, 从 n =8 开始, 然后是 n =6, n =4, 最后是 n =2, 这时我们得到

$$I_8 = \frac{7}{8} I_6 = \frac{7}{8} \cdot \frac{5}{6} I_4 = \frac{7}{8} \cdot \frac{5}{6} \cdot \frac{3}{4} I_2 = \frac{7}{8} \cdot \frac{5}{6} \cdot \frac{3}{4} \cdot \frac{1}{2} I_0.$$

现在需要计算 I_0. 因为 $\cos^0$ 的结果是 1, 所以 $I_0 = \int_0^{\pi/2} 1\mathrm{d}x = \pi/2$. 化简上述分式, 我们得到

$$\int_0^{\pi/2} \cos^8(x)\mathrm{d}x = \frac{7 \cdot 5 \cdot 3 \cdot 1}{8 \cdot 6 \cdot 4 \cdot 2} \times \frac{\pi}{2} = \frac{35\pi}{256}.$$

作为我们辛苦计算的奖励, 我们可以容易地计算出 $\int_0^{\pi/2} \cos^n(x)\mathrm{d}x$ 的值 (对于任何正整数 n). (为了计算奇次幂的值, 有必要知道 $I_1 = \int_0^{\pi/2} \cos(x)\mathrm{d}x = 1$.)

顺便说一下, 约化公式不必使用三角恒等式. 例如, 计算

$$I_n = \int x^n \mathrm{e}^x \mathrm{d}x,$$

你需要用到分部积分法, 通过设 $u = x^n, \mathrm{d}v = \mathrm{e}^x \mathrm{d}x$ (所以有 $v = \mathrm{e}^x$) 去计算

$$I_n = x^n \mathrm{e}^x - \int n x^{n-1} \mathrm{e}^x \mathrm{d}x.$$

这样有了约化公式 $I_n = x^n \mathrm{e}^x - n I_{n-1}$. 顺便说一下, 这次我们是用 I_{n-1} 来表示 I_n, 而不像前几个三角函数的例子, 用 I_{n-2} 来表示 I_n. 所以在这个链式的最后你仅仅需要知道 I_0, 不难发现 $I_0 = \int \mathrm{e}^x \mathrm{d}x = \mathrm{e}^x + C$.

19.3 关于三角换元法的积分

现在, 让我们看看怎样计算关于二次函数平方根的奇次幂的积分. 一些典型的例子有

$$\int \frac{\mathrm{d}x}{x^3\sqrt{x^2-4}} \quad 或 \quad \int \frac{x^2}{(9-x^2)^{3/2}}\mathrm{d}x \quad 或 \quad \int (x^2+15)^{-5/2}\mathrm{d}x.$$

基本思想是: 有三种情况, 分别为 $a^2 - x^2$、$x^2 + a^2$、$x^2 - a^2$, 这里 a 为常数. 例如上面的第一个积分是当 a =2 时 $x^2 - a^2$ 的情况, 第二个积分是当 a =3 时 $a^2 - x^2$ 的情况, 第三个积分是当 $a = \sqrt{15}$ 时 $x^2 + a^2$ 的情况. 这三种情况要求不同的换元法. 大多数的情况下, 在换元之后, 都会得到一个关于三角函数的幂的被积函数, 这正是我们在前几节见过的. 下面我们一次研究一种情况, 最后再做个总结.

19.3.1 类型 1: $\sqrt{a^2 - x^2}$

如果你遇到关于 $\sqrt{a^2 - x^2}$ 的奇次幂的积分, 正确的换元是使用 $x = a\sin(\theta)$. 如果你喜欢也可以使用 $x = a\cos(\theta)$, 但这并没有任何优势, 所以我们依然使用正弦函数, 因为这个替代很有效果:

$$a^2 - x^2 = a^2 - a^2\sin^2(\theta) = a^2(1-\sin^2(\theta)) = a^2\cos^2(\theta),$$

现在可以容易地求平方根. 请记住, 如果你把变量从 x 改到 θ, 那么就该由从以 x 为变量转到以 θ 为变量的积分. 也就是说, 积分符号里的每一个 x 都要用 θ 来表示. 具体地, 我们需要用带有 θ 的变量以及 $\mathrm{d}\theta$ 表示 $\mathrm{d}x$. 没问题, 仅仅需要对方程 $x = a\sin(\theta)$ 求微分就可以得到 $\mathrm{d}x = a\cos(\theta)\mathrm{d}\theta$. (这种类型的替代在 18.1.2 节和 18.1.3 节中讨论过, 但当时是以 x 为变量而不是以这个替代变量求解.) 无论如何, 现在积分是以 θ 为变量了, 但在最后还需要再换回到以 x 为变量的积分. 为此, 我们要画一个锐角是 θ 的直角三角形, 这会很有帮助 (如图 19-1 所示).

我们知道 $\sin(\theta) = x/a$, 所以可以设出这两个边的边长 (如图 19-2 所示).

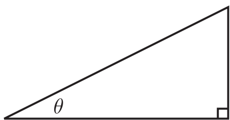


图 19-1

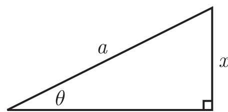


图 19-2

最后使用毕达哥拉斯定理可得第三边边长为 $\sqrt{a^2-x^2}$, 这个三角形就能确定了 (如图 19-3 所示).
现在, 我们使用这个三角形可以很容易地计算出 $\cos(\theta)$、$\tan(\theta)$ 或其他任何关于 θ 的三角函数的值, 也能方便地转换回到以 x 为变量的积分.

来看看怎样实际应用的. 我们使用刚才的例子：

$$\int \frac{x^2}{(9-x^2)^{3/2}}\mathrm{d}x.$$

我们设 $x = 3\sin(\theta)$, 以此完成替代, 所以 $\mathrm{d}x = 3\cos(\theta)\mathrm{d}\theta$. 同时我们也看到 $9 - x^2 = 9 - 9\sin^2(\theta) = 9\cos^2(\theta)$. 所以这个积分为

$$\int \frac{(3\sin(\theta))^2}{(9\cos^2(\theta))^{3/2}} \cdot 3\cos(\theta)\mathrm{d}\theta = \frac{3^2\times 3}{9^{3/2}}\int \frac{\sin^2(\theta)}{\cos^3(\theta)}\cos(\theta)\mathrm{d}\theta = \int \tan^2(\theta)\mathrm{d}\theta,$$

因为 $9^{3/2} = 27$. 使用 19.2.2 节中的方法可得

$$\int \tan^2(\theta)\mathrm{d}\theta = \int (\sec^2(\theta) - 1)\mathrm{d}\theta = \tan(\theta) - \theta + C.$$

现在我们需要做的是换回到以 x 为变量的状态. 因为 $\sin(\theta) = x/3$, 这个相关的三角形如图 19-4 所示.

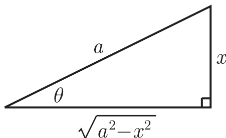


图 19-3

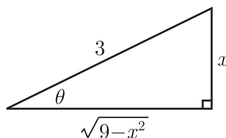


图 19-4

从这个三角形中可得 $\tan(\theta)=x/\sqrt{9-x^2}$. 同时, 因为 $\sin(\theta)=x/3$, 我们有 $\theta=\sin^{-1}(x/3)$. 把这些换回到答案中, 我们得到

$$\int\frac{x^2}{(9-x^2)^{3/2}}\mathrm{d}x=\frac{x}{\sqrt{9-x^2}}-\sin^{-1}\left(\frac{x}{3}\right)+C.$$

如果不使用三角形, 你可能会把 $\tan(\theta)$ 写为烦琐的形式:

$$\tan\left(\sin^{-1}\left(\frac{x}{3}\right)\right),$$

但我希望你能更认可我们得到的答案.

讨论类型 2 之前, 你发现我们在这里有些大意了吗? 需要计算出 $(9\cos^2(\theta))^{3/2}$, 但仅仅说它等于 $27\cos^3(\theta)$. 当然, $9^{3/2}=27$, 但这就能说明 $(\cos^2(\theta))^{3/2}=\cos^3(\theta)$ 吗? 实际上仅仅当 $\cos(\theta)\geqslant 0$ 时才成立. 问题是对一个数值求它的 3/2 次幂, 实际上是要求这个数值的平方根. 对于任何正数 A, 我们有 $A^{3/2}=(A^{1/2})^3=(\sqrt{A})^3$. 所以应该写为

$$(\cos^2(\theta))^{3/2}=(\sqrt{\cos^2(\theta)})^3=|\cos^3(\theta)|.$$

幸运的是, 这个绝对值对于类型 1 和类型 2 没有必要 (但对于类型 3 就不是这样了), 所以我们所做的一切是正确的. 这个观点将会在 19.3.6 节中详细讨论.

19.3.2　类型 2: $\sqrt{x^2+a^2}$

如果一个积分是关于 $\sqrt{x^2+a^2}$ 的奇次幂, 那么正确的换元是 $x=a\tan(\theta)$. 这种方法很有效果, 因为

$$x^2+a^2=a^2\tan^2(\theta)+a^2=a^2(\tan^2(\theta)+1)=a^2\sec^2(\theta).$$

并且我们需要知道 $\mathrm{d}x=a\sec^2(\theta)\mathrm{d}\theta$. 因为 $\tan\theta=x/a$, 所以这个三角形如图 19-5 所示.

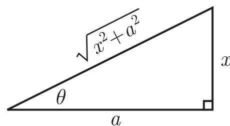


图　19-5

现在我们来看这个例子:

$$\int(x^2+15)^{-5/2}\mathrm{d}x.$$

这里使用换元法, 设 $x=\sqrt{15}\tan(\theta)$. 我们有 $\mathrm{d}x=\sqrt{15}\sec^2(\theta)\mathrm{d}\theta$, 并且注意 $x^2+15=15\tan^2(\theta)+15=15\sec^2(\theta)$. 这个积分变为

$$\begin{aligned}\int(15\sec^2(\theta))^{-5/2}\sqrt{15}\sec^2(\theta)\mathrm{d}\theta&=\frac{15^{1/2}}{15^{5/2}}\int(\sec(\theta))^{-5}\sec^2(\theta)\mathrm{d}\theta\\&=(15)^{-2}\int\cos^3(\theta)\mathrm{d}\theta.\end{aligned}$$

(我们再一次做了一件有风险的事情: 用 $15^{-5/2}\sec^{-5}(\theta)$ 替代 $\left(15\sec^2(\theta)\right)^{-5/2}$, 完全忽略了绝对值符号. (如果你提前阅读了 19.3.6 节, 就会知道其中的原因了.) 我们仍然需要计算 $15^{-2}\int\cos^3(\theta)\mathrm{d}\theta$. 让我们使用 19.2.1 节中的方法. 请注意被积函数是 $\cos(\theta)$ 的奇次幂, 所以我们可以提出一个 $\cos(\theta)$ 项, 用 $\sin(\theta)$ 做换元:

$$(15)^{-2}\int\cos^3(\theta)\mathrm{d}\theta=(15)^{-2}\int(1-\sin^2(\theta))\cos(\theta)\mathrm{d}\theta$$
$$=(15)^{-2}\left(\sin(\theta)-\frac{\sin^3(\theta)}{3}\right)+C.$$

(我在这里忽略了换元的细节, 因为我确信你能自己做出这道题.) 现在, 回到以 x 为变量的状态. 因为 $\tan(\theta)=x/\sqrt{15}$, 由此得到如图 19-6 所示的三角形.

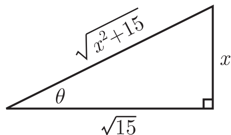


图 19-6

从这个三角形中, 你能简单地发现 $\sin(\theta)=x/\sqrt{x^2+15}$, 也就是说

$$\int(x^2+15)^{-5/2}\mathrm{d}x=(15)^{-2}\left(\sin(\theta)-\frac{\sin^3(\theta)}{3}\right)+C$$
$$=\frac{1}{225}\left(\frac{x}{\sqrt{x^2+15}}-\frac{x^3}{3(x^2+15)^{3/2}}\right)+C.$$

(你知道为什么 $\sin^3(\theta)=x^3/(x^2+15)^{3/2}$ 吗? 仅仅把里面的 $\sin(\theta)$ 用 $x/(x^2+15)^{\frac{1}{2}}$ 替代即可.)

19.3.3 类型 3: $\sqrt{x^2-a^2}$

最后, 关于 $\sqrt{x^2-a^2}$ 的奇次幂的情况又怎样呢? 正确的换元是 $x=a\sec(\theta)$, 因为

$$x^2-a^2=a^2\sec^2(\theta)-a^2=a^2(\sec^2(\theta)-1)=a^2\tan^2(\theta),$$

你能容易地得到平方根. 为了做这个换元, 我们需要知道 $\mathrm{d}x=a\sec(\theta)\tan(\theta)\mathrm{d}\theta$. 因为 $\sec(\theta)=x/a$, 这个三角形如图 19-7 所示.

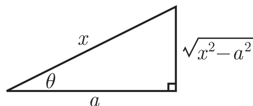


图 19-7



例如, 计算

$$\int\frac{\mathrm{d}x}{x^3\sqrt{x^2-4}},$$

设 $x=2\sec(\theta)$, 所以有 $\mathrm{d}x=2\sec(\theta)\tan(\theta)\mathrm{d}\theta, x^2-4=4\tan^2(\theta)$, 这个积分变为

$$\int\frac{2\sec(\theta)\tan(\theta)}{(2\sec(\theta))^3\sqrt{4\tan^2(\theta)}}\mathrm{d}\theta=\int\frac{2\sec(\theta)\tan(\theta)}{8\sec^3(\theta)\times 2\tan(\theta)}\mathrm{d}\theta$$
$$=\frac{1}{8}\int\frac{1}{\sec^2(\theta)}\mathrm{d}\theta=\frac{1}{8}\int\cos^2(\theta)\mathrm{d}\theta.$$

如果这次用 $2\tan(\theta)$ 替代 $\sqrt{4\tan^2(\theta)}$, 那就大错特错了. 像我们将要在 19.3.6 节中看到的那样, 它只有在 $x>0$ 的时候才是正确的. 所以让我们做个假设. 我们需要计算 $\frac{1}{8}\int\cos^2(\theta)\mathrm{d}\theta$. 因为 cos 是偶次幂, 所以要使用 19.2.1 节的倍角公式:

$$\frac{1}{8}\int\cos^2(\theta)\mathrm{d}\theta=\frac{1}{8}\int\frac{1}{2}(1+\cos(2\theta))\mathrm{d}\theta=\frac{\theta}{16}+\frac{\sin(2\theta)}{32}+C.$$

很好, 我们只需要再回到以 x 为变量的状态. 这里需要点小技巧, 让我们使用三角形 (如图 19-8 所示) 来帮助计算.

问题是我们需要知道 $\sin(2\theta)$ 的值. 为了计算这个数值, 要使用三角公式：

$$\sin(2\theta)=2\sin(\theta)\cos(\theta).$$

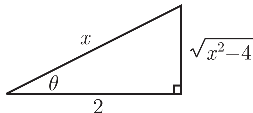


图　19-8

由此可知 $\sin(\theta)=\sqrt{x^2-4}/x, \cos(\theta)=2/x$, 再把它们带回到原结果中, 可得

$$\begin{aligned}\int\frac{\mathrm{d}x}{x^3\sqrt{x^2-4}}&=\frac{1}{16}\sec^{-1}\left(\frac{x}{2}\right)+\frac{1}{32}\cdot 2\cdot\frac{\sqrt{x^2-4}}{x}\cdot\frac{2}{x}+C\\&=\frac{1}{16}\sec^{-1}\left(\frac{x}{2}\right)+\frac{\sqrt{x^2-4}}{8x^2}+C.\end{aligned}$$

请记住, 它仅当 $x>0$ 时才成立. 我们将在 19.3.6 节中重视这个例子, 并考虑 $x\leqslant 0$ 的情况.

19.3.4　配方和三角换元法

在我们总结这种方法之前还有一点需要说明. 有时, 你可能需要求解关于 $\sqrt{\pm x^2+ax+b}$ 的奇次幂的积分. 也就是说, 你有了一次项 ax, 这样情况就复杂了. 求解这个积分的方法很简单：我们可以配方, 然后做替代得到刚才介绍的三种情况. 例如, 计算

$$\int(x^2-4x+19)^{-5/2}\mathrm{d}x,$$

首先配方 (参见 1.6 节的配方方法)：

$$x^2-4x+19=(x^2-4x+4)-4+19=(x-2)^2+15.$$

所以, 要计算的积分实际上是

$$\int((x-2)^2+15)^{-5/2}\mathrm{d}x.$$

设 $t=x-2$, 所以 $\mathrm{d}t$=$\mathrm{d}x$, 那么这是一个以 t 为变量的积分了：

$$\int(t^2+15)^{-5/2}\mathrm{d}t,$$

这就得到了一个在 19.3.2 节中已经计算过的积分! 该题目的答案是 (此时以 t 为变量)：

$$\frac{1}{225}\left(\frac{t}{\sqrt{t^2+15}}-\frac{t^3}{3(t^2+15)^{3/2}}\right)+C,$$

用 $x-2$ 替代 t, 得到

$$\int(x^2-4x+19)^{-5/2}\mathrm{d}x=\frac{1}{225}\left(\frac{x-2}{\sqrt{x^2-4x+19}}-\frac{(x-2)^3}{3(x^2-4x+19)^{3/2}}\right)+C.$$

这种方法的准则是, 带有一次项的二次函数可以通过配方再换元的方法求得结果.

19.3.5 关于三角换元法的总结

让我们用一个表格来总结刚才使用过的针对三种类型积分的换元法:

类型 1: $\sqrt{a^2-x^2}$	类型 2: $\sqrt{x^2+a^2}$	类型 3: $\sqrt{x^2-a^2}$
设 $x=a\sin(\theta)$ $\mathrm{d}x=a\cos(\theta)\mathrm{d}\theta$ $a^2-x^2=a^2\cos^2(\theta)$	设 $x=a\tan(\theta)$ $\mathrm{d}x=a\sec^2(\theta)\mathrm{d}\theta$ $x^2+a^2=a^2\sec^2(\theta)$	设 $x=a\sec(\theta)$ $\mathrm{d}x=a\sec(\theta)\tan(\theta)\mathrm{d}\theta$ $x^2-a^2=a^2\tan^2(\theta)$
a, x, θ, $\sqrt{a^2-x^2}$	$\sqrt{x^2+a^2}$, x, θ, a	x, $\sqrt{x^2-a^2}$, θ, a

下一节将要介绍遇到 $a^2\cos^2(\theta)$ 和 $a^2\tan^2(\theta)$ 情况时, 什么时候 (及为什么) 可以去掉绝对值符号. 你第一次遇到这种情况时, 可能会忽略它, 但之后又不得不重新考虑.

19.3.6 平方根的方法和三角换元法

我们以前提过, 这一节可能会有些烦琐. 你还跟得上吗? 很好, 现在我们回到类型 1. 我们直接把 $\sqrt{a^2\cos^2(\theta)}$ 化简为 $a\cos(\theta)$, 完全忽略了 $\cos(\theta)$ 的绝对值. 实际上, 我们写 $x=a\sin(\theta)$ 时, 是在说 $\theta=\sin^{-1}(x/a)$.

但 θ 在哪里呢? 很好, 从 10.2.1 节中, 我们知道 $\sin^{-1}$ 的范围是 $[-\pi/2,\pi/2]$; 也就是说, θ 在第一或第四象限, 所以 $\cos(\theta)$ 一直都是非负的. 在此, 我们不需要任何绝对值符号!

类型 2 也是同样的. 在这种情况下, 我们把 $\sqrt{a^2\sec^2(\theta)}$ 化简为 $a\sec(\theta)$. 可以不使用绝对值吗? 我们设 $x=a\tan(\theta)$, 所以 $\theta=\tan^{-1}(x/a)$. 因为 $\tan^{-1}$ 的值域是 $(-\pi/2,\ \pi/2)$, 所以这次的 θ 也在第一或第四象限. 就是说, $\sec(\theta)$ 一直都是正的, 所以此次我们也不需要绝对值符号.

但对于类型 3, 我们就不这么走运了. 这次我们需要化简 $\sqrt{a^2\tan^2(\theta)}$, 但它的结果不一定为 $a\tan(\theta)$. 你看, 因为 $x=a\sec(\theta)$, 我们有 $\theta=\sec^{-1}(x/a)$. 如果你看看 10.2.4 节, 会发现 $\sec^{-1}$ 的值域是 $[0,\pi]$, 但不包括 $\pi/2$ 这一点. 所以 θ 在一二象限, $\tan(\theta)$ 既可能为正也可能为负. 但至少它同 x 有着同样的符号, 可以通过 $y=\sec^{-1}(x)$ 的图像来判断.

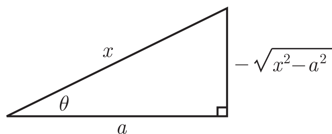


图 19-9

所以当 $x>0$ 时, 我们认为 $\sqrt{a^2\tan^2(\theta)}=a\tan(\theta)$. 另一方面, 当 $x<0$ 时, 我们需要写为 $-a\tan(\theta)$. 在这种情况下, 三角形如图 19-9 所示.

一个三角形有两条边是负的 (分别是 x 和 $-\sqrt{x^2-a^2}$), 这确实有些怪异, 但这却便于我们记忆, 因为这个三角函数的所有符号都是正确的. 在 19.3.3 节的例子

$$\int \frac{\mathrm{d}x}{x^3\sqrt{x^2-4}}$$

中, 我们知道当 $x>0$ 时, 这个积分的结果为

$$\frac{1}{16}\sec^{-1}\left(\frac{x}{2}\right)+\frac{\sqrt{x^2-4}}{8x^2}+C.$$

(当 $x>0$ 时, x 实际上要大于 2, 否则分子中的 $\sqrt{x^2-4}$ 项就失去意义了.) 现在让我们重新计算当 $x<0$ 时的情况. 我们仍然设 $x=2\sec(\theta)$, 但是现在要用 $-2\tan(\theta)$ 替代 $\sqrt{4\tan^2(\theta)}$. 与之前唯一的不同就是负号:

$$\begin{aligned}\int \frac{\mathrm{d}x}{x^3\sqrt{x^2-4}} &= \int \frac{2\sec(\theta)\tan(\theta)}{(2\sec(\theta))^3\sqrt{4\tan^2(\theta)}}\mathrm{d}\theta\\ &= \int \frac{2\sec(\theta)\tan(\theta)}{8\sec^3(\theta)\times(-2\tan(\theta))}\mathrm{d}\theta\\ &= -\frac{1}{8}\int \cos^2(\theta)\mathrm{d}\theta = -\frac{\theta}{16}-\frac{2\sin(\theta)\cos(\theta)}{32}+C.\end{aligned}$$

我们回到以 x 为变量的状态, 需要使用一个修正的三角形 (如图 19-10 所示).

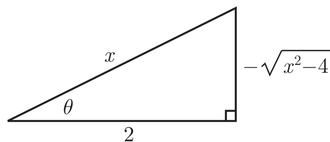


图　19-10

因此, 实际上 $\sin(\theta)=-\sqrt{x^2-4}/x$, $\cos(\theta)=2/x$. 注意 $\sin(\theta)$ 实际上是大于零的, 因为 $x<0$. 现在再带回到原积分可得到

$$\begin{aligned}\int \frac{\mathrm{d}x}{x^3\sqrt{x^2-4}} &= -\frac{1}{16}\sec^{-1}\left(\frac{x}{2}\right)-\frac{1}{32}\cdot 2\cdot\frac{-\sqrt{x^2-4}}{x}\cdot\frac{2}{x}+C\\ &= -\frac{1}{16}\sec^{-1}\left(\frac{x}{2}\right)+\frac{\sqrt{x^2-4}}{8x^2}+C.\end{aligned}$$

这就是当 $x<0$ 时的答案, 同我们刚才的答案几乎是一样的, 只是 sec 的反函数需要一个负号. 当然, 常数 C 同 $x>0$ 时的 C 是不同的. 为什么呢? 因为我们正在寻找一个函数使它的导数为 $1/x^3\sqrt{x^2-4}$, 它的定义域为 $(-\infty,-2)\cup(2,\infty)$. 所以, 它的反导数实际上分为两部分, 每一部分可由另一部分上下平移而得. 总而言之, 完整的答案是:

$$\int \frac{\mathrm{d}x}{x^3\sqrt{x^2-4}} = \begin{cases}\dfrac{1}{16}\sec^{-1}\left(\dfrac{x}{2}\right)+\dfrac{\sqrt{x^2-4}}{8x^2}+C_1 & \text{当 } x>2 \text{ 时},\\ -\dfrac{1}{16}\sec^{-1}\left(\dfrac{x}{2}\right)+\dfrac{\sqrt{x^2-4}}{8x^2}+C_2 & \text{当 } x<-2 \text{ 时}.\end{cases}$$

其中 C_1 和 C_2 是不同的. 其实我们遇到过这样的积分, 例如 $\int 1/x\mathrm{d}x$, 它的积分结果里就有两个常数. 参见 17.7 节. 在实际应用中, 遇到类型 3 的问题时, 我们常常

只考虑 $x>0$ 时的情况. 这样可以避免上述烦琐的情况, 并且取平方根不用担心符号. 但当 $x<0$ 时, 你需要注意更多细节.

19.4 积分技巧总结



我们已经介绍了很多计算积分的方法. 问题是, 对于一道计算积分的题应该使用哪种方法呢? 有时这不容易, 你可能在发现正确方法之前要试很多种不同的方法, 甚至需要把多种方法混合在一起. 下面是一些帮助你解决问题的技巧.

- 当你看到题目时, 会发现一种显而易见的换元, 那就试试它. 例如, 被积函数中的一部分是另一部分的导数, 那么就使用 t 做换元.
- 如果 $\sqrt[n]{ax+b}$ 这种形式出现在被积函数中, 就像在 18.1.2 节那样, 设 $t=\sqrt[n]{ax+b}$.
- 对于有理函数的积分 (也就是说, 两个多项式的商), 看分子是否为分母导数的倍数. 如果是, 可以通过设 “$t=$ 分母” 来计算. 另外, 也可以使用部分分式法 (参见 18.3 节).
- 若观察后没有发现明显的换元可用, 可使用这一章介绍的方法:
  - 关于 $\sqrt{1-\cos(x)}$ 或 $\sqrt{1+\cos(x)}$ 的函数, 使用倍角公式;
  - 关于 $1-\sin^2(x)$、$1-\cos^2(x)$、$1+\tan^2(x)$、$\sec^2(x)-1$、$\csc^2(x)-1$ 或 $1+\cot^2(x)$ 的函数, 使用毕达哥拉斯恒等式: $\sin^2(x)+\cos^2(x)=1$、$\tan^2(x)+1=\sec^2(x)$ 或 $1+\cot^2(x)=\csc^2(x)$;
  - 关于 $1\pm\sin(x)$ (或与其相似的情况) 在分母时的函数, 分子分母同时乘以它的共轭表达式, 然后试着使用毕达哥拉斯定理;
  - 关于 $\cos(mx)\cos(nx)$、$\sin(mx)\sin(nx)$ 或 $\sin(mx)\cos(nx)$ 的函数的积分, 使用积化和差公式;
  - 关于三角函数的次幂的积分, 应该学会从 19.2.1 节到 19.2.5 节的所有方法.
- 如果被积函数是关于 $\sqrt{x^2-a^2}$ 这种形式的奇次幂的情况 (例如 $(x^2-a^2)^{3/2}$, $(x^2-a^2)^{5/2}$ 等), 或 $\sqrt{x^2+a^2}$ 或 $\sqrt{a^2-x^2}$ 等类似情况的奇次幂形式, 那么使用三角换元法 (但要先校验是否有明显的换元). 如果二次函数包含一次项, 那么先配方. 更多细节参见 19.3 节.
- 如果被积函数是乘积的形式, 同时也没有明显的换元可用, 那么可以考虑分部积分法. (参见 18.2 节.)
- 如果没有可用的换元法, 被积函数又是 $\ln(x)$ 的幂或反三角函数的形式, 那么可以考虑使用分部积分法. 在这种情况下, 设 u 是 $\ln(x)$ 的幂或为适当的反三角函数. 例如, 计算

$$\int \frac{\ln(1+x^2)}{x^2}\mathrm{d}x.$$

首先校验没有换元法可用; 因为没有任何灵感, 所以我们用分部积分法. 等一下, 它不是乘积的形式! 再等一下, 商也可以写成乘积的形式! 让我们把它重写为

$$\int \ln(1+x^2) \times \frac{1}{x^2}\mathrm{d}x,$$



这时再用分部积分法, 设 $u = \ln(1+x^2), \mathrm{d}v = (1/x^2)\mathrm{d}x$. 现在试试, 你会得到答案

$$-\frac{\ln(1+x^2)}{x} + 2\tan^{-1}(x) + C.$$

即使你掌握了所有的方法, 如果你不做大量的练习, 那么遇到实际问题时, 还是会陷入混乱. 在做了大量的练习后, 你就能应付各种各样复杂的积分, 能够在计算中找到自信. 这样, 你就是一个优秀的积分计算者了.

第 20 章 反常积分：基本概念

这个主题比较难, 我分两章来讨论. 本章介绍反常积分, 下一章给出更详细的讨论, 介绍怎样解决关于反常积分的一些问题. 如果你是第一次阅读本章, 那么应该读明白这里的每一个知识点. 但如果你正在备考, 我建议你忽略本章, 但请注意方框内的公式和标记为重要的部分, 集中精力看下一章. 下面是我们在这一章将要学习的内容：

- 反常积分、收敛和发散的定义;
- 关于没有边界区域的反常积分;
- 关于比较判别法、极限比较判别法、p 判别法和绝对收敛判别法的理论基础.

在下一章中, 我们还会再次介绍这四种判别法, 并给出一些应用示例.

20.1 收敛和发散

到底什么是反常积分？在第 16 章, 我们见过积分

$$\int_a^b f(x)\mathrm{d}x.$$

该被积函数 f 如果在 $[a, b]$ 区间内是有界的, 并是连续的 (如有有限个间断点也可), 那么这个积分就是有意义的. 如果这个积分有无限多个不连续点, 该积分也可能是有意义的 (参见 16.7 节中的例子). 但如果函数 f 不是有界的, 情况又怎样呢？这就是说, 当 x 在区间 $[a, b]$ 内时, 函数 f 的值越来越大 (正方向或负方向, 或两个方向). 当函数 f 在这个区间有一条垂直渐近线时会出现这种情况：函数在渐近线附近变得很大, 且没有界限. 这就使上述积分成了反常积分.

即使函数 f 是有界的, 也会出现一种不同类型的无界. 这个闭区间 $[a, b]$ 实际上是无界的, 如 $[0,\ \infty)$、$[-7,\ \infty)$、$(-\infty,\ 3]$, 甚至 $(-\infty,\ \infty)$. 这也使这个积分成为反常积分.

所以, 如果出现下面的情况, 积分 $\int_a^b f(x)\mathrm{d}x$ 就是反常积分：

(1) 函数 f 在闭区间 $[a,b]$ 内是无界的;

(2) $b=\infty$;

(3) $a=-\infty$.

从现在开始, 我们集中精力研究第一种情况, 在随后的 20.2 节再研究后两种情况. 像我以前说过的那样, 如果一个函数在某个位置有垂直渐近线, 那么该函数在

这个位置是无界的, 尽管可能会有一些奇怪的走势. (例如函数 $f(x)=\frac{1}{x}\sin\left(\frac{1}{x}\right)$, 当 x 趋于 0 时, 它的图像是大幅振荡的.) 如果函数 $f(x)$ 在 x 接近于某点 c 时是无界的, 那么我们说该函数在 $x=c$ 点有一个破裂点. 大多数情况下, 它就是指有垂直渐近线.

所以, 我们来看看函数在 $x=a$ 点有垂直渐近线时的简单情况, 如图 20-1 所示.

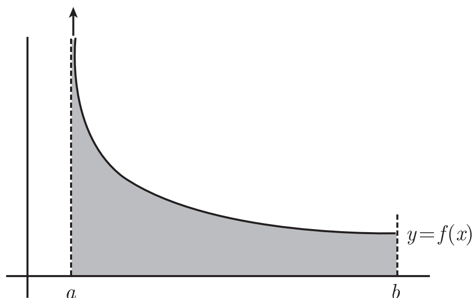


图　20-1

如果我说积分 $\int_a^b f(x)\mathrm{d}x$ 是上图中阴影部分的面积 (平方单位), 那么我是在说谎. 问题是, 由于是垂直渐近线该区域会一直延伸到这页的最上部且还会一直延伸下去, 该区域越来越狭长.

由于该区域不停地向上延伸, 那么其面积就是无限的. 这个结论是正确的吧? 不一定. 如果该区域足够狭长, 那会出现一个数学奇迹, 面积就是有限的了. 为了研究什么情况下一块无限区域的面积会是有限的, 我们需要使用极限. 其基本思想是: 设 ε 是一个很小的正数, 函数 f 在区间 $[a+\varepsilon,b]$ 上是可积的, 因为函数 f 在此是有界的. 你会得到一些有限的数. 现在, 用一个更小的数 ε 去重复这种情况, 你会得到一个新的有限的数, 如图 20-2 所示.

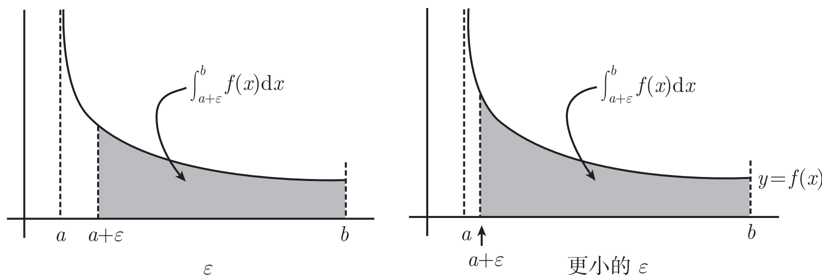


图　20-2

数 ε 越小, 我们对这块无限区域的估算就越接近于真实值. 这说明, 我们应该用越来越小的 ε 重复这个过程, 当 $\varepsilon\to 0^+$ 时, 看看我们是否会得到一个极限 L. 如果可以, 我们就把 L 解释为正在计算的这块区域的面积. 在这种情况下, 我们说积

分 $\int_a^b f(x)\mathrm{d}x$收敛于数 L. 如果没有极限, 我们就不能对于这块区域找到一个有意义的答案, 只能放弃寻找, 认为该积分是发散的. 注意：**如果积分不是反常积分, 那么它自然收敛!** 在实践中, 只要这个函数是有界的且区间 $[a,b]$ 是有界的, 那么就可以说这样的积分是收敛的, 因为它甚至不反常. 它仅仅是一些有限的数.

现在, 当在 $x=a$ 处有破裂点时, 我们有：

> 如果仅仅在 x 接近于 a 点该函数 $f(x)$ 是无界的, 则定义
> $$\int_a^b f(x)\mathrm{d}x = \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\mathrm{d}x.$$

在此假设这个极限存在. 如果能找到这样的极限, 我们就说这个积分收敛; 否则认为该积分发散. 就像其他的极限一样, 由于极限的结果可能为 ∞ 或 $-\infty$, 或当 $\varepsilon\to 0^+$ 时它的图像上下振荡, 因而这个极限可能没有意义.

这让我们认识到非常重要的一点. 当看到一个反常积分时, 我们需要知道的一件非常重要的事情是：它是收敛的还是发散的. 这个积分的收敛值是多少并不是很重要 (假设它是收敛的). 在实际中, 如果你知道积分是收敛的, 可以通过复杂的计算求得收敛值. 如果积分是发散的, 而你使用计算机估算这个积分值, 那么可能会得到意想不到的结果. 计算机还不能真正理解无限或疯狂的上下振荡.

20.1.1 反常积分的一些例子

考虑两个积分：

$$\int_0^1 \frac{1}{x}\mathrm{d}x \quad 和 \quad \int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x.$$

这两个都是反常积分, 因为它们的被积函数在 $x=0$ 点都有垂直渐近线. 所以我们可以使用刚才方框里的公式. 在第一种情况下, 我们有

$$\int_0^1 \frac{1}{x}\mathrm{d}x = \lim_{\varepsilon\to 0^+}\int_\varepsilon^1 \frac{1}{x}\mathrm{d}x = \lim_{\varepsilon\to 0^+} \ln|x|\Big|_\varepsilon^1 = \lim_{\varepsilon\to 0^+}(\ln(1)-\ln(\varepsilon)) = \infty.$$

(我们使用了这些性质：$\ln(1)=0$; 当 $\varepsilon\to 0^+$ 时 $\ln(\varepsilon)\to-\infty$.) 因为得到的答案是正无穷, 所以这个反常积分是发散的. 第二个积分的情况又如何呢？再次使用公式, 我们有

$$\int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x = \lim_{\varepsilon\to 0^+}\int_\varepsilon^1 \frac{1}{x^{1/2}}\mathrm{d}x = \lim_{\varepsilon\to 0^+} 2x^{1/2}\Big|_\varepsilon^1 = \lim_{\varepsilon\to 0^+}(2\sqrt{1}-2\sqrt{\varepsilon}) = 2.$$

我们得到了一个有限的数, 所以这个积分是收敛的. 我们已经证明了这个积分收敛于 2, 但如上节最后所述, 我们并不在意收敛值, 主要研究它是收敛的还是发散的, 而不是收敛于多少.

我们得到了什么？为什么反常积分 $\int_0^1 1\,/\,x\mathrm{d}x$ 是发散的, 而 $\int_0^1 1\,/\,\sqrt{x}\mathrm{d}x$ 却是

收敛的? 毕竟, 两个积分的图像 $y = 1/x$ 和 $y = 1/\sqrt{x}$ 是非常相似的 (如图 20-3 所示).

当然, 这两个被积函数是不同的. 当 $0 < x < 1$ 时, $1/x$ 比 $1/\sqrt{x}$ 要大. 从几何角度来解释, $1/\sqrt{x}$ 的图像实际上比 $1/x$ 的图像更接近于 y 轴. 可以说, $1/\sqrt{x}$ 的图像足够接近于 y 轴, 以至于它所对应的积分是收敛的; 而 $1/x$ 没有那么接近于 y 轴, 所以它所对应的积分是发散的. 但糟糕的是, 对于所有在 $x = 0$ 点有渐近线的函数, 很难区分哪个函数足够接近于 y 轴, 哪个足够远离于 y 轴. 大多数情况下, 你需要分别对待每个积分.

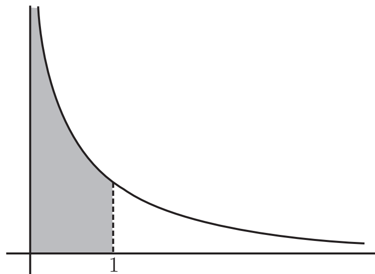


图 20-3

这里有一点非常重要. 假设你遇到反常积分 $\int_a^b f(x)\mathrm{d}x$, 它的被积函数只在 $x = a$ 点有垂直渐近线, 你仅仅想知道该积分是收敛的还是发散的. 这时, b 的值对我们没有影响, 可以把它换成任意大于 a 的有限的数, 只要不选择新的垂直渐近线或新的破裂点. 之所以这样说, 首先请看 (根据定义):

$$\int_a^b f(x)\mathrm{d}x = \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^b f(x)\mathrm{d}x,$$

在此假设这个极限存在. 现在让我们把 b 换成其他任意的数 c, 但要比数 a 大. 如果 $x = a$ 依然是函数 f 的破裂点, 那么我们有

$$\int_a^c f(x)\mathrm{d}x = \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^c f(x)\mathrm{d}x,$$

再次假设极限是存在的. 我们可以在 $x = b$ 点把最后一个积分分开 (16.3 节介绍过这种方法) 而得到

$$\int_a^c f(x)\mathrm{d}x = \lim_{\varepsilon \to 0^+} \left(\int_{a+\varepsilon}^b f(x)\mathrm{d}x + \int_b^c f(x)\mathrm{d}x \right).$$

ε 对第二个积分没有任何影响; 实际上, 因为函数 f 在 b 和 c 点之间是有界的, 所以这个积分收敛于一个数 M. 我们已经证明了

$$\int_a^c f(x)\mathrm{d}x = \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^b f(x)\mathrm{d}x + M.$$

如果右侧的极限存在, 那么积分 $\int_a^b f(x)\mathrm{d}x$ 收敛. 增加一个数 M 它仍然为有限的, 所以 $\int_a^b f(x)\mathrm{d}x$ 还是收敛的. 相反, 如果极限不存在, 这时增加一个数 M 对情况也没有影响, 所以 $\int_a^b f(x)\mathrm{d}x$ 和 $\int_a^c f(x)\mathrm{d}x$ 同时发散.

我们已经证明了, 一个反常积分在有界区间的收敛或发散是由它的被积函数在

非常接近破裂点时的走势决定的. 在具体情况下, 因为我们知道积分 $\int_0^1 1 \;/\; x\mathrm{d}x$ 是发散的, 所以可以得出

$$\int_0^2 \frac{1}{x}\mathrm{d}x, \quad \int_0^{100} \frac{1}{x}\mathrm{d}x, \quad \int_0^{0.000\,000\,1} \frac{1}{x}\mathrm{d}x$$

这些积分都是发散的. 另一方面, 因为 $\int_0^1 1 \;/\; \sqrt{x}\mathrm{d}x$ 是收敛的, 我们也可以说

$$\int_0^2 \frac{1}{\sqrt{x}}\mathrm{d}x, \quad \int_0^{100} \frac{1}{\sqrt{x}}\mathrm{d}x, \quad \int_0^{0.000\,000\,1} \frac{1}{\sqrt{x}}\mathrm{d}x$$

这些积分是收敛的. 所有的这些都是发生在 $x = 0$ 的垂直渐近线附近.

20.1.2 其他破裂点

对于积分 $\int_a^b f(x)\mathrm{d}x$, 如果函数 f 仅仅在积分上限 b (而不是 a) 是无界的, 那么我们可以应用刚才的方法. 仅仅的不同是, 我们这次需要从左方而不是右方趋于 b. 所以

如果函数仅仅在 x 接近于 b 点是无界的, 则定义

$$\int_a^b f(x)\mathrm{d}x = \lim_{\varepsilon\to 0^+}\int_a^{b-\varepsilon} f(x)\mathrm{d}x,$$

在此假设这个极限存在; 如果它不存在, 那么像前面的例子一样, 它是发散的.

如果函数 f 在区间 $[a, b]$ 内有破裂点 c, 那该怎么办呢? 在这种情况下, 如果函数 f 仅仅是在该区间 (a, b) 内的 c 点无界, 那么需要把这个积分分成两部分：

$$\int_a^c f(x)\mathrm{d}x \quad 和 \quad \int_c^b f(x)\mathrm{d}x.$$

实际上, 我们知道怎样使用极限定义这些积分 —— 使用方框里的公式, 上述积分分别为

$$\lim_{\varepsilon\to 0^+}\int_a^{c-\varepsilon} f(x)\mathrm{d}x \quad 和 \quad \lim_{\varepsilon\to 0^+}\int_{c+\varepsilon}^b f(x)\mathrm{d}x,$$

关键点是：只有当这两部分的积分都收敛时, 积分 $\int_a^b f(x)\mathrm{d}x$ 才是收敛的; 如果任何一个发散, 那么整个积分就是发散的. 毕竟, 你不能把一个不存在的东西加到另一个上去. 无论另一个是否存在, 都不能这样做.



这个例子激发了我们第一个灵感：为计算反常积分, 如果必要就把它分解. 每一部分最多只能有一个瑕点, 而且该点要在积分的上下限上. (在这里, 瑕点指的是破裂点, 下一节我们会看到一个不同的 "瑕点", 它不是 "破裂点".)

来看积分

$$I = \int_0^3 \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x,$$



这个被积函数的瑕点是 $x = 0, 1, 2$ 和 -1. 最后这个点对我们的计算没有影响, 因为积分区间仅在 0 到 3 之间. 另外三个却很重要. 我们需要在这些瑕点之间选择一些数比如 1/2 和 3/2, 因为它们对计算不会产生影响. 现在, 我们把原始的积分分成下面 5 个积分:

$$I_1 = \int_0^{1/2} \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x, \quad I_2 = \int_{1/2}^{1} \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x,$$

$$I_3 = \int_1^{3/2} \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x, \quad I_4 = \int_{3/2}^{2} \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x,$$

$$I_5 = \int_2^{3} \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x.$$

注意, 这 5 个积分的瑕点都不超过一个, 且这些点都在积分的上下限位置. 积分 I_1、I_3 和 I_5 的瑕点在积分的下限, 而 I_2 和 I_4 的瑕点在积分的上限. 原始积分收敛的唯一可能性是从 I_1 到 I_5 都是收敛的. 如果它们都是收敛的, 那么积分 I 的值是从 I_1 到 I_5 的和. (实际上, 这 5 个积分没有一个是收敛的! 在 21.5 节, 我们将会看到为什么这样.)

20.2　关于无穷区间上的积分

现在要研究当积分上下限有一个或同是无穷时的情况; 也就是说, 积分区间是无界的. 为计算

$$\int_a^{\infty} f(x)\mathrm{d}x,$$

其中 a 是常数, 函数 f 在区间 $[a, \infty)$ 没有破裂点, 我们需要使用另一个极限方法. 这次, 我们对 $[a, N]$ 区间求积分, 其中 N 是个很大的数. 这会给出一个非常好的值且是有限的. 随着 N 值的增大, 我们重复这个过程, 会得到不同的计算结果. 继续计算, 看看最后的积分结果到底是多少. 如果极限确实存在, 那么这个积分是收敛的; 否则它就是发散的. 用符号来表示, 我们可以定义

$$\boxed{\int_a^{\infty} f(x)\mathrm{d}x = \lim_{N\to\infty}\int_a^{N} f(x)\mathrm{d}x,}$$

在此假设这个极限是存在的. 在这种情况下, 这个积分是收敛的; 否则它是发散的. 同 20.1.1 节中最后描述的原因一样, a 的值与计算结果无关. 所以只要你不选择函数 f 的新的破裂点, 那么 a 的值对广义积分的收敛还是发散就没有任何影响. 仅仅需要考虑的是当 x 非常大时, 函数 $f(x)$ 的走势怎样.

如果在区间 $(-\infty, b]$ 上函数 f 没有其他的破裂点, 我们可以使用相似的定义

$$\boxed{\int_{-\infty}^{b} f(x)\mathrm{d}x = \lim_{N\to\infty}\int_{-N}^{b} f(x)\mathrm{d}x.}$$

如果函数 f 在整个区间内都没有破裂点, 那么要计算

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x$$

该是怎样的呢？

尽管它没有破裂点, 但仍然有两个瑕点：∞ 和 $-\infty$. 是的, 每当出现 ∞ 和 $-\infty$ 时, 我们都把它们看作瑕点, 因为需要分别对待它们. 我们把上述积分分成两部分, 这样每一部分只有一个瑕点. 选一个你喜欢的数 (我选 0), 考虑这两个积分：

$$\int_{-\infty}^{0} f(x)\mathrm{d}x \quad 和 \quad \int_{0}^{\infty} f(x)\mathrm{d}x.$$

我们知道这两个积分分别意味着什么, 当然只有两个积分都收敛时原始积分才收敛. 如果你选了一个不同于 0 的数, 对计算结果也没有任何影响, 因为积分的收敛或发散不是由端点值决定的.



下面是一些关于无界区间上的积分例子. 考虑积分

$$\int_{1}^{\infty} \frac{1}{x}\mathrm{d}x \quad 和 \quad \int_{1}^{\infty} \frac{1}{x^2}\mathrm{d}x.$$

第一个是

$$\lim_{N\to\infty}\int_{1}^{N} \frac{1}{x}\mathrm{d}x = \lim_{N\to\infty} \ln|x|\Big|_1^N = \lim_{N\to\infty}(\ln(N)-\ln(1)) = \infty,$$

而第二个是

$$\lim_{N\to\infty}\int_{1}^{N} \frac{1}{x^2}\mathrm{d}x = \lim_{N\to\infty} -\frac{1}{x}\Big|_1^N = \lim_{N\to\infty}\left(-\frac{1}{N}+1\right) = 1.$$

所以, 第一个积分是发散的, 第二个积分是收敛的.

这里有一个问题：积分

$$\int_{0}^{\infty} \frac{1}{x}\mathrm{d}x \quad 和 \quad \int_{0}^{\infty} \frac{1}{x^2}\mathrm{d}x$$

是收敛的还是发散的？因为这两个积分的瑕点是 0 和 ∞, 所以需要把它们分成两部分. 第一个可以分解为

$$\int_{0}^{1} \frac{1}{x}\mathrm{d}x \quad 和 \quad \int_{1}^{\infty} \frac{1}{x}\mathrm{d}x.$$

注意是否选择 1 作为分界点由你来决定. 选什么并不重要 (只要它是一个正数)! 无论如何, 我们已经看到这两个积分都是发散的, 所以积分 $\int_0^\infty 1/x\mathrm{d}x$ 也是发散的.

我们把第二个积分也分成两部分：

$$\int_0^1 \frac{1}{x^2}\mathrm{d}x \quad 和 \quad \int_1^\infty \frac{1}{x^2}\mathrm{d}x.$$

我们已经看到, 它是收敛的. 对于第一个积分, 我们可以使用关于极限的公式, 但也还有一种不是很显见的方法. 其基本思想是：我们已经看到 $\int_0^1 1/x\mathrm{d}x$ 发散到无穷大. 如果你仔细考虑就会发现, 当 x 在 0 和 1 之间时, $1/x^2$ 比 $1/x$ 大. (对吗？因为在区间 (0, 1) 之间, x^2 比 x 小, 所以它们倒数的大小正好相反.) 所以, 如果在 [0, 1] 区间 $1/x$ 下的面积是无限的, 那么在该区间内函数 $1/x^2$ 下的面积会更大, 所以它也是无限的! 不需要做任何其他工作, 我们就可得出结论：$\int_0^1 1/x^2\mathrm{d}x$ 也是发散的. 于是整个积分 $\int_0^\infty 1/x^2\mathrm{d}x$ 是发散的, 但它发散的原因不是因为积分上限的无穷大, 而是因为积分下限的 0. 注意我们比较 $1/x^2$ 和 $1/x$ 的方法是比较判别法, 在下一节要研究它.

20.3　比较判别法 (理论)

假设有两个非负函数, 它们至少在某些区间上是非负的. 如果第一个函数比第二个函数大, 第二个函数的积分 (在这个区间内) 是发散的, 那么第一个函数的积分 (在同样的区间内) 也是发散的. 从数学角度上可以这样来解释. 我们想知道积分 $\int_a^b f(x)\mathrm{d}x$ 的情况, 但现在仅仅知道积分 $\int_a^b g(x)\mathrm{d}x$ 的情况. 如果在区间 (a,b) 内, 函数 $f(x) \geqslant g(x) \geqslant 0$, 且积分 $\int_a^b g(x)\mathrm{d}x$ 是发散的, 那么积分 $\int_a^b f(x)\mathrm{d}x$ 也是发散的. 事实上, 因为 $f(x) \geqslant g(x)$, 所以可以写为

$$\int_a^b f(x)\mathrm{d}x \geqslant \int_a^b g(x)\mathrm{d}x = \infty.$$

因而第一个积分也是发散的. 在上述例子中, 我们只需写

$$\int_0^1 \frac{1}{x^2}\mathrm{d}x \geqslant \int_0^1 \frac{1}{x}\mathrm{d}x = \infty,$$

并知道不等式左侧是发散的. 当然, 我们已经知道右侧也是发散的.

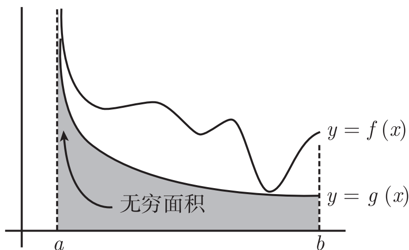


图　20-4

当看图 20-4 所示的图像时, 我们就会更清楚这种情况了.

在这个图像中, 在 $x = a$ 和 $x = b$ 之间的 $y = g(x)$ 的面积被认为是无穷的. 函数 $y = f(x)$ 的图像在函数 $y = g(x)$ 的上方, 所以它的面积 (在 $x = a$ 和 $x = b$ 之间) 应该更大. 比无穷还大当然仍是无穷大. 所以积分 $\int_a^b f(x)\mathrm{d}x$ 也是发散的.

如果 $f(x) \leqslant g(x)$, 积分 $\int_a^b g(x)\mathrm{d}x$ 仍然是发散的, 情况又会是怎样的呢？你会对积分 $\int_a^b f(x)\mathrm{d}x$ 得出怎样的结论呢？答案是：两种情况都有可能. 也就是说什么

结论都得不到. 我们从数学角度来解释这个问题：

$$\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x = \infty$$

所以正在求的积分 $\int_a^b f(x)\mathrm{d}x$ 小于或等于无穷大. 也就是说, 它可能是小于无穷的, 所以是收敛的; 它也可能是等于无穷的, 所以是发散的. 很好, 我们知道它既可能是收敛的也可能是发散的. 我们没有得到任何结论, 所以这个条件什么都没有给我们.

另一方面, 对于收敛性, 方向要反过来. 是这样的：我们想知道积分 $\int_a^b f(x)\mathrm{d}x$ 的情况, 但现在知道积分 $\int_a^b g(x)\mathrm{d}x$ 是收敛的, 那么我们希望 $f(x) \leqslant g(x)$. 你可能会说, 我们希望函数 f 是由函数 g 控制的. 很好, 这时我们已经可以确定收敛性了 (仍然假设两个函数都是正的). 也就是说, 如果在区间 $(a,\ b)$ 内 $0 \leqslant f(x) \leqslant g(x)$, 且积分 $\int_a^b g(x)\mathrm{d}x$ 是收敛的, 那么积分 $\int_a^b f(x)\mathrm{d}x$ 也一定是收敛的. 数学上的表示形式是

$$\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x < \infty,$$

所以两个积分都是收敛的 (注意左边的积分是正的, 所以它不可能发散到 $-\infty$), 见图 20-5.

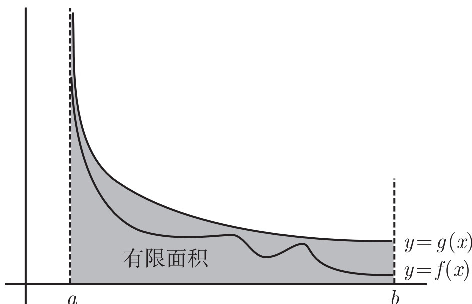


图 20-5

对于 $x=a$ 和 $x=b$ 之间的 $y=g(x)$ 的阴影部分面积, 我们假设它是有限的. 你可以清楚地看到, 所要研究的面积, 即函数 $y=f(x)$ 在 $x=a$ 和 $x=b$ 之间的面积, 比有限的阴影面积要小. 因为我们想要的面积是正的并小于一个有限的面积, 所以它是有限的.

请注意：假设你知道积分 $\int_a^b g(x)\mathrm{d}x$ 是收敛的, 但你有个相反的不等式 $f(x) \geqslant g(x)$. 现在你想要分析的图像 $y=f(x)$ 在另一条曲线 $y=g(x)$ 的上方. 这很不好, 我们只能得到

$$\int_a^b f(x)\mathrm{d}x \geqslant \int_a^b g(x)\mathrm{d}x.$$

所以, 不等式左侧的积分大于或等于一个有限的数, 积分可能是有限的也可能是无

限的. 这相当于没有得到任何结论. 我们又白费力气了!



到目前为止, 从数学角度看, 我们还没有正式说明比较判别法. 实际上, 这种方法并不是很复杂. 把积分分解开是必要的, 我们已经了解了其基本思想. 例如, 如果函数 f 和 g 在 $x=a$ 点都有垂直渐近线, 在其他地方没有破裂点, 且区间 $[a,b]$ 内的所有 x 都有 $0 \leqslant f(x) \leqslant g(x)$, 那么我们有

$$0 \leqslant \int_{a+\varepsilon}^{b} f(x)\mathrm{d}x \leqslant \int_{a+\varepsilon}^{b} g(x)\mathrm{d}x$$

对于任何 $\varepsilon > 0$ 成立. 现在取极限. 如果反常积分 $\int_a^b g(x)\mathrm{d}x$ 收敛, 那么不等式右边就是个有限的数. 现在的情况由中间的那个积分来决定. 因为函数 $f(x)$ 一直都为正, 所以当 ε 趋于 0 时, 这个中间积分变得越来越大. 虽然如此, 但它再大也大不过积分 $\int_a^b g(x)\mathrm{d}x$, 而这个积分恰恰就是一个有限的数. 所以唯一的可能性是: 当 $\varepsilon \to 0^+$ 时, 这个中间积分收敛于一个有限的数[①]. 简而言之, 积分 $\int_a^b f(x)\mathrm{d}x$ 是收敛的. 这样, 我们从收敛角度 (上述的第二个反常积分) 证明了比较判别法, 在上述特殊情况下, 函数 f 和 g 仅仅在 $x=a$ 点出现瑕点. 我把证明发散的部分留给你去做, 而且也要说明在 $x=b$ 点出现瑕点的情况. 这同证明收敛没什么不同. 当然, 如果瑕点出现在积分的中间, 或有多个瑕点, 在使用比较判别法之前就需要把积分分成几个部分.

在下一章中, 我们将会看到应用比较判别法的更多例子. 现在我们去看看另一个判别法.

20.4　极限比较判别法 (理论)

比较判别法是用一个函数的反常积分的结果去判别另一个函数的反常积分. 极限比较判别法是类似的, 但并不需要一个比被判别的函数更大的函数. 相反, 我们仅仅需要两个近似的函数. 其基本思想是: 假设有两个函数在破裂点 $x=a$ 是非常接近的 (它们再也没有其他的破裂点), 那么积分 $\int_a^b f(x)\mathrm{d}x$ 和 $\int_a^b g(x)\mathrm{d}x$ 同时收敛或同时发散, 它们的行为是相同的. 直观上讲, 这个说法是行得通的, 我们来仔细说说什么叫两个函数是“非常接近”的.

20.4.1　函数互为渐近线

假设有两个函数 f 和 g 满足

$$\lim_{x\to a} \frac{f(x)}{g(x)} = 1.$$

这就是说, 当 x 接近于 a 时, $f(x)/g(x)$ 的比值是非常接近于 1 的. 如果比值是 1, 那么函数 $f(x)$ 和 $g(x)$ 是相等的; 因为比值仅仅是接近于 1, 所以 $f(x)$ 是非常接近

① 事实上, 这个显而易见的陈述非常重要, 正是它将 $\mathbb{R}$ 与 $\mathbb{R}$ 的任一包含所有有理数的真子集区分开来.

于 $g(x)$ 的. 但这并不意味着函数 $f(x)$ 和 $g(x)$ 的差是非常小的! 例如, (对于同样的值 x) 函数 $f(x)$ 可能是万亿, 而 $g(x)$ 可能是万亿加上一百万; 在这种情况下, 比值 $f(x)/g(x)$ 比 1 略小, 而 $f(x)$ 和 $g(x)$ 的差却是一百万! 但从另一个角度说, 这两个数是非常接近的, 因为它们之间的差一百万相对于它们自己的数值是非常小的.

所以, 我们说如果比值的极限是 1, 那么当 $x \to a$ 时, $f(x) \sim g(x)$; 即

当 $x \to a$ 时, $f(x) \sim g(x)$ 同 $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = 1$ 有着同样的意义.

这并不是说明当 x 接近于 a 时, $f(x)$ 大约等于 $g(x)$; 它说明当 x 接近于 a 时, $f(x)$ 和 $g(x)$ 的比值接近于 1. 我们说当 $x \to a$, 函数 $f(x)$ 和 $g(x)$ 是渐近等价的. 当然你可以用 $x \to \infty$ 或 $x \to a^+$ 来替代 $x \to a$, 只需要在极限中做同样的替代.

所有这些都可能是无用的, 除非我们有这样形式的极限：

$$\lim_{x\to a} \frac{f(x)}{g(x)} = 1.$$



实际上, 我们已经见过很多这种形式的极限! 这有些例子[1]：

$$\lim_{x\to\infty} \frac{3x^3 - 1000x^2 + 5x - 7}{3x^3} = 1, \quad \lim_{x\to 0} \frac{\sin(x)}{x} = 1,$$

$$\lim_{x\to 0} \frac{\mathrm{e}^x - 1}{x} = 1, \quad \lim_{x\to 0} \frac{\ln(1+x)}{x} = 1.$$

第一个极限可以写为：当 $x \to \infty$ 时 $3x^3 - 1000x^2 + 5x - 7 \sim 3x^3$. 也就是说, 当 $x \to \infty$ 时, $3x^3 - 1000x^2 + 5x - 7$ 和 $3x^3$ 是渐近等价的. 同理, 第二个极限表明, 当 $x \to 0$ 时, $\sin(x) \sim x$. 第三个和第四个极限表明, 当 $x \to 0$ 时, $\mathrm{e}^x - 1$ 和 $\ln(1+x)$ 同 x 是渐近等价的; 也就是说, 当 $x \to 0$, $\mathrm{e}^x - 1 \sim x$ 和 $\ln(1+x) \sim x$.

我们只是以不同的形式重写了每一个极限, 但这是一种很方便的形式. 实际上, 你可以对渐近等价的函数做幂运算, 然后得到一对新的渐近等价的函数. 例如, 我们知道当 $x \to 0$ 时有 $\sin(x) \sim x$, 则可以立刻写出, 当 $x \to 0$ 时有 $\sin^3(x) \sim x^3$, 或者 $1/\sin(x) \sim 1/x$. 你也可以用其他像 x 一样趋于 0 的量替代 x, 比如 x 的幂. 例如, 从 $x \to 0$ 时$\sin(x) \sim x$ 开始, 我们用 $4x^7$ 替代 x, 可看到当 $x \to 0$ 时, $\sin(4x^7) \sim 4x^7$. 你甚至可以让两个渐近等价的函数相除或相乘, 假设它们的极限对应的 x 值相同. 例如, 我们知道当 $x \to 0$ 时 $\tan(x) \sim x$, 因为

$$\lim_{x\to 0} \frac{\tan(x)}{x} = 1.$$

所以我们能把 $\sin(x) \sim x$ 和 $\tan(x) \sim x$ 乘到一起, 得到当 $x \to 0$ 时的渐近关系 $\tan(x)\sin(x) \sim x^2$.

加或减这些关系却不适用上述规则. 例如当 $x \to 0$ 时, 以 $\tan(x) \sim x$ 和 $\sin(x) \sim$

[1] 这些例子可以分别在 4.3 节、7.1.1 节、9.4.2 节和 9.4.3 节中找到.


x 开始, 那么不能从第一个中减去第二个得到 $\tan(x) - \sin(x) \sim x - x$. $x - x$ 是 0, 没有什么能同 0 是渐近等价的. 为什么没有呢? 因为, 如果当 $x \to a$ 时有 $f(x) \sim 0$, 这时我们有

$$\lim_{x \to a} \frac{f(x)}{0} = 1.$$

这讲不通, 因为等式的左边没有任何意义. 所以, 可以对这种渐近等价关系做乘积、除法、取幂, 但一定不要做加法和减法.

20.4.2　关于判别法的陈述

好了, 现在我们已经有了渐近等价两个函数的概念, 也有了一些例子 (如 $x \to 0$ 时 $\sin(x) \sim x$). 那又怎样呢? 假设某个函数 f, 它的瑕点仅仅在 a 点, 你想知道反常积分 $\int_a^b f(x)\mathrm{d}x$ 是收敛还是发散的. 如果当 x 趋近于 a 时, 你能找到一个函数 g 的走势非常接近于 f, 那么可以用函数 g 替代函数 f, 判断积分 $\int_a^b g(x)\mathrm{d}x$ 是收敛的还是发散的. 无论你得到什么关于 g 的结论都适用于 f.

更正式地说, 如果当 $x \to a$ 时 $f(x) \sim g(x)$, 且这两个函数在区间 $[a, b]$ 上没有其他的瑕点了, 那么积分 $\int_a^b f(x)\mathrm{d}x$ 和 $\int_a^b g(x)\mathrm{d}x$ 是同时收敛或同时发散的. (如果同时收敛, 它们的收敛值可能不同.) 这就是*极限比较判别法*. 这只是粗略的介绍, 我们将在下一章给出更多的例子. 假设我们想知道积分



$$\int_0^1 \frac{1}{\sin(\sqrt{x})}\mathrm{d}x$$

是收敛的还是发散的. 看起来求解 $1/\sin\sqrt{x}$ 的反导数不是一件容易的事情. 很幸运, 我们不需要求它的反导. 因为当 $x \to 0$ 时$\sin(x) \sim x$, 所以可以用一个更小的量 $\sqrt{x}$ 替代这个很小的量 x, 这样可得当 $x \to 0^+$ 时 $\sin(\sqrt{x}) \sim \sqrt{x}$. (我们需要使用 $x \to 0^+$, 因为仅仅当 $x \geqslant 0$ 时, $\sqrt{x}$ 才有意义.) 两边同时取倒数, 可得

$$\text{当 } x \to 0^+ \text{ 时,} \quad \frac{1}{\sin(\sqrt{x})} \sim \frac{1}{\sqrt{x}}.$$

请注意在区间 (0, 1] 上, $1/\sin\sqrt{x}$ 和 $1/\sqrt{x}$ 没有破裂点. 所以极限比较判别法告诉我们, 积分

$$\int_0^1 \frac{1}{\sin(\sqrt{x})}\mathrm{d}x \quad \text{和} \quad \int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x$$


同时收敛或同时发散. 我们用一个简单的积分 $\int_0^1 1/\sqrt{x}\mathrm{d}x$ 替代了一个较难的积分. 从 20.1.1 节中, 我们已经知道这个简单的积分是收敛的, 所以立刻知道要计算的积分 (左边那个) 也是收敛的.

当然有些判别法也适用于破裂点在 b 点或积分区间是无界的情况. 我们将在 21.2 节列举所有的情况. 现在, 让我们看看为什么这个判别法适用于上述例子. 因为当 $x \to a$ 时 $f(x) \sim g(x)$, 所以我们知道

$$\lim_{x\to a}\frac{f(x)}{g(x)}=1.$$

特别地, 假设足够趋于 a, 那么比值 $f(x)/g(x)$ 至少是 1/2 且不比 2 大. 也就是说, 我们能在 a 和 b 的区间内选一个数 c, 满足

$$\frac{1}{2}\leqslant\frac{f(x)}{g(x)}\leqslant 2,\quad x\in(a,c].$$

这个不等式可以重写为

$$\frac{1}{2}g(x)\leqslant f(x)\leqslant 2g(x), x\in(a,c].$$

现在就能使用比较判别法了. 例如, 如果积分 $\int_a^b g(x)\mathrm{d}x$ 是发散的, 那么积分 $\int_a^c g(x)\mathrm{d}x$ 也是 (像我们已经见过的那样). 事实上, $\frac{1}{2}\int_a^c g(x)\mathrm{d}x$ 也是发散的, 直觉解释就是, 无穷的一半还是无穷! 所以, 函数 $f(x)$ 比 $\frac{1}{2}g(x)$ 大说明积分 $\int_a^c f(x)\mathrm{d}x$ 是发散的, 说明 $\int_a^b f(x)\mathrm{d}x$ 也是发散的. 另一方面, 如果积分 $\int_a^b g(x)\mathrm{d}x$ 是收敛的, 那么积分 $2\int_a^c g(x)\mathrm{d}x$ 也是收敛的, 我们能再次使用比较判别法证明积分 $\int_a^b f(x)\mathrm{d}x$ 也是收敛的 (你自己可以证明一下).



附注：大多数教材关于极限比较判别法都有不同的陈述. 特别地, $f(x)/g(x)$ 的极限可能实际上不是 1 —— 可能是任何正数, 这时上述陈述 (稍微修正之后) 依然成立. 另一方面, 极限不是 1 并没有什么意义, 且无法使用直观的 $\sim$ 表示法. 在下一章, 我们将能非常熟练地使用这个判别法.

20.5 p 判别法 (理论)

我们有了比较判别法和极限比较判别法, 需要知道怎样去使用它们. 我们的基本策略是 (下一章将细致讲解)：选择一个能与函数 f 相比较的函数 g. 我们希望函数 g 足够简单到可以判断它是收敛的还是发散的.

问题是, 我们能选择什么样的 g 函数? 最常用的函数是 $1/x^p$, 其中 $p>0$. 我们已经看到一些关于 $1/x$、$1/\sqrt{x}$ 和 $1/x^2$ 的积分, 它们分别对应于 $p=1$、$\frac{1}{2}$ 和 2. 因为这些函数很容易求得积分, 所以可以使用极限公式得到 p 判别法.

- **p 判别法, $\int^\infty$ 的情况：** 对于任何有限值 $a>0$, 积分

$$\int_a^\infty \frac{1}{x^p}\mathrm{d}x$$

在 $p>1$ 时是收敛的, 在 $p\leqslant 1$ 时是发散的.

- **p 判别法, $\int_0$ 的情况：** 对于任何有限值 $a>0$, 积分

$$\int_0^a \frac{1}{x^p}\mathrm{d}x$$

在 $p<1$ 时是收敛的, 在 $p\geqslant 1$ 时是发散的.

注意, 这两种情况是相反的, 只是 $p=1$ 除外. 其中的积分

$$\int_0^a \frac{1}{x^p}\mathrm{d}x \quad \text{或} \quad \int_a^\infty \frac{1}{x^p}\mathrm{d}x$$

是收敛的, 而另一个积分是发散的. $p=1$ 的情况对应于 $1/x$, 我们已经见过, 这两个积分在这种情况下都是发散的.

p 判别法真的很有用, 其实际应用很广泛, 所以千万不要混淆这两种情况! 要记住这种方法的正确情况, 就要记住 $1/x^2$ 和 $1/\sqrt{x}$ 的情况. 我仅仅记得两个事实:

$$\int_a^\infty \frac{1}{x^2}\mathrm{d}x \text{ 收敛}, \int_0^a \frac{1}{\sqrt{x}}\mathrm{d}x \text{ 也收敛}.$$

依据这两个事实, 我就能记住整个 p 判别法! 怎样记住的? 由第一种情况我们知道, 趋于 ∞ 时的情况与趋于 0 时的情况是相反的, 我知道积分

$$\int_0^a \frac{1}{x^2}\mathrm{d}x$$

是发散的; 同理, 由第二种情况我们知道, 积分

$$\int_a^\infty \frac{1}{\sqrt{x}}\mathrm{d}x$$

也是发散的. 那么其他指数的情况是什么样子呢? 任何高于 1 的指数 (例如 3/2、2 或 70) 的趋势同 $1/x^2$ 是一样的, 而任何低于 1 的指数 (例如 1/2、2/3 或 0.999) 的趋势同 $1/\sqrt{x}$ 是一样的 (记住, 它就是 $1/x^{1/2}$).

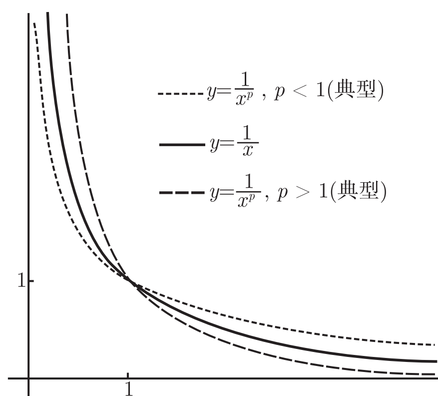


图　20-6

观察图 20-6 会很有帮助. 在这个图像中, 点状线和虚线是典型的 $p<1$ 或 $p>1$ 时 $y=1/x^p$ 的图像. 实线是 $y=1/x$ 的图像, 它不足够接近于 y 轴使积分 $\int_0^1 1/x\mathrm{d}x$ 收敛, 也不足够接近于 x 轴使 $\int_1^\infty 1/x\mathrm{d}x$ 收敛. 另一方面, 对于任何 $p<1$, $\int_0^1 1/x^p\mathrm{d}x$ 是收敛的, 因为点状线足够接近于 y 轴. 当你查看 x 轴时, 这种情况是相反的: 这时需要查看虚线, 即 $p>1$ 时的 $y=1/x^p$, 它足够接近于 x 轴, 所以 $\int_1^\infty 1/x^p\mathrm{d}x$ 是收敛的.

注意, 因为 $1.000\,000\,1>1$, 所以积分

$$\int_1^\infty \frac{1}{x^{1.000\,000\,1}}\mathrm{d}x$$

收敛, 尽管 $\int_1^\infty 1/x\mathrm{d}x$ 是发散的! 仅仅把 x 的幂从 1 变为 1.000 000 1, 这个微小的变化足够引起质变了. 这展示出了收敛和发散的精妙之处.

现在我们证明 p 判别法. 幸运的是, 这仅仅需要使用 20.1 节的公式. 首先, 考虑积分



$$\int_a^\infty \frac{1}{x^p}\mathrm{d}x,$$

其中常数 $a>0$. 如果 $p=1$, 该积分变为 $1/x$, 我们已经知道, 这个积分在这种情况下是发散的. 另外我们有

$$\int_a^\infty \frac{1}{x^p}\mathrm{d}x = \lim_{N\to\infty}\int_a^N x^{-p}\mathrm{d}x = \lim_{N\to\infty}\frac{1}{1-p}x^{1-p}\Big|_a^N$$
$$= \frac{1}{1-p}\left(\left(\lim_{N\to\infty}N^{1-p}\right)-a^{1-p}\right).$$

现在如果极限

$$\lim_{N\to\infty}N^{1-p}$$

存在, 那么整个积分 $\int_a^\infty 1/x^p\mathrm{d}x$ 是收敛的. 如果这个极限不存在, 那么该积分是发散的. 所以, 将上述极限重写为

$$\lim_{N\to\infty}\frac{1}{N^{p-1}}.$$

如果 $p>1$, 那么 $p-1>0$, 所以当 N 很大时 N^{p-1} 非常大; 它的倒数变得很小, 所以极限是 0, 最初的积分是收敛的. 另一方面, 如果 $p<1$, 那么 $p-1<0$, 所以 N^{1-p} 非常大, 这个极限趋于无穷大, 说明原始积分是发散的. 这证明了 p 判别法的一半. 另一半证明与此类似, 只是使用 $\varepsilon\to 0^+$ 而不是 $N\to\infty$. 我将证明细节留给你.



20.6 绝对收敛判别法

比较判别法的一个假设是函数 f 和 g 都是非负的. 但如果你想判断一个负函数的走势, 该怎么办呢? 如果这个函数一直为负, 则可把负号提出来后把它归为正函数的情况. 在下一章中, 我们将会看到例子. 另一方面, 如果这个函数在积分区间不停地在正负之间振荡, 则可以应用**绝对收敛判别法**. 陈述如下:

如果 $\int_a^b |f(x)|\mathrm{d}x$ 是收敛的, 那么 $\int_a^b f(x)\mathrm{d}x$ 也是收敛的.

这对于无限区间上的积分也是适用的 (例如 $[a,\infty)$ 而不是 $[a,b]$). 注意: 如果原始积分的绝对值是发散的, 那么这个原始积分可能还是收敛的! 这样的例子很酷, 但超过了本书的范围. 我们在 23.7 节讨论正交级数时将看到一些相似情况.



为什么上述方法是有用的? 首先, $|f(x)|$ 总是非负的, 所以可以使用反常积分的比较判别法. 例如, 考虑反常积分



$$\int_1^\infty \frac{\sin(x)}{x^2}\mathrm{d}x.$$

当 x 越来越大时, 被积函数 $\frac{\sin(x)}{x^2}$ 在正负之间振荡, 所以不能使用比较判别法[①]或极限比较判别法. 让我们先试着使用绝对收敛判别法.

我们需要考虑积分

$$\int_1^\infty \left|\frac{\sin(x)}{x^2}\right|\mathrm{d}x.$$

这可以重写为

$$\int_1^\infty \frac{|\sin(x)|}{x^2}\mathrm{d}x,$$

因为 x^2 不可能为负. 现在就能使用比较判别法了. 因为对于所有的 x, $|\sin(x)| \leqslant 1$, 所以对于所有的 x, 有

$$\frac{|\sin(x)|}{x^2} \leqslant \frac{1}{x^2}.$$

比较判别法表明

$$\int_1^\infty \frac{|\sin(x)|}{x^2}\mathrm{d}x \leqslant \int_1^\infty \frac{1}{x^2}\mathrm{d}x.$$

因为根据 p 判别法, 不等式的右侧是收敛的, 所以左侧的积分也是收敛的. 最后, 我们使用绝对收敛判别法得

$$\int_1^\infty \frac{|\sin(x)|}{x^2}\mathrm{d}x\text{收敛, 所以}\int_1^\infty \frac{\sin(x)}{x^2}\mathrm{d}x\text{也收敛.}$$

这里有一点比较微妙, 真的需要使用绝对值.

还有一个例子:

$$\int_0^\infty \cos(x)\mathrm{d}x.$$

这个被积函数 $\cos(x)$ 在正负之间振荡, 所以应该先研究它的绝对值情况:

$$\int_0^\infty |\cos(x)|\mathrm{d}x.$$

遗憾的是, 这个新的积分不可能是收敛的. 想要知道为什么吗? 画一个 $y = |\cos(x)|$ 的图像, 你会看到大量相似的小山丘, 一个接一个. 要把这些无限个小山丘加到一起得到有限的值是不可能的. 所以这个绝对值型的积分是发散的. 这说明我们不能使用绝对收敛判别法! 只有当该积分的绝对值情况是收敛的, 才能使用这个方法.

我们需要重新来过. 我们不知道最原始的积分是收敛还是发散的. 所以, 我们使用瑕点在 ∞ 的反常积分的定义:

---

① 直接比较是不行的, 因为积分 $\int_1^N \sin(x)/x^2\mathrm{d}x$ 并不随 N 变大而变大. 20.3 节结尾处的结论不适用, 因为它需要在不达到 $g(x)$ 的积分上限时 $f(x)$ 的积分越来越大.

$$\int_0^\infty \cos(x)\mathrm{d}x = \lim_{N\to\infty}\int_0^N \cos(x)\mathrm{d}x = \lim_{N\to\infty}\sin(x)\Big|_0^N$$
$$= \lim_{N\to\infty}(\sin(N)-\sin(0)) = \lim_{N\to\infty}\sin(N).$$

最后一个极限并不存在, 因为 $\sin(N)$ 在 -1 到 1 之间反复振荡, 即使 N 一直无限变大也如此. 所以, 原始积分 $\int_0^\infty \cos(x)$ 发散的原因是振荡太多, 而不是因为它趋于 ∞ 或 $-\infty$.

振荡的积分处理起来很复杂. 如果你足够幸运, 可以像上面一样使用标准定义. 大多数情况下, 这根本不起作用. 很多数学家花费了大量的时间想弄明白这一点. 此刻, 只要将上例记在心间就可以了, 我们在下一章还有很多事情要处理, 届时会再次讨论判别法, 并着重解决反常积分相关的问题.



在此之前, 我们简单看一下为什么绝对收敛判别法可以起作用. 假设我们知道

$$\int_a^b |f(x)|\mathrm{d}x$$

收敛. 有一个很好的技巧：设对于 f 的 x 的定义区间 $[a,b]$, 有 $g(x)=|f(x)|+f(x)$. 那么 g 有两个重要特性：首先, $g(x)\geqslant 0$; 其次, $g(x)\leqslant 2|f(x)|$. (两种情况下, 都假定 x 是 f 定义区间 $[a,b]$ 中的任意数.) 实际上, 稍考虑一下就可以知道, 每当 $f(x)\geqslant 0$ 都有 $g(x)=2f(x)$, 而且每当 $f(x)<0$ 都有 $g(x)=0$. 尝试证明, 从这点可得到前面提到的 g 的两个重要特性.



无论如何, 我们现在可以对 g 使用比较判别法了：

$$0\leqslant \int_a^b g(x)\mathrm{d}x \leqslant 2\int_a^b |f(x)|\mathrm{d}x < \infty.$$

结论是

$$\int_a^b g(x)\mathrm{d}x$$

也是收敛的. 那又怎样呢？注意, $f(x)=g(x)-|f(x)|$, 所以

$$\int_a^b f(x)\mathrm{d}x = \int_a^b g(x)\mathrm{d}x - \int_a^b |f(x)|\mathrm{d}x.$$

右侧的两个积分都是收敛的 —— 第一个已经证明了, 第二个假设为收敛的 —— 所以, 左侧的积分也是收敛的.

第 21 章 反常积分：如何解题

我们实际应用一下, 并看看有关反常积分的例子. 讨论过程中, 我们会总结主要的方法. 在上一章中, 我们介绍了一些很有用的判别方法. 为了更加有效地利用这些方法, 你需要了解一些常见函数的性质, 特别是它们在 0 和 ∞ 附近是怎样变化的. 所谓常见函数是指：多项式函数、三角函数、指数函数和对数函数. 下面是这一章要讲的内容：

- 首次遇到反常积分时需要做什么, 包括怎么处理被积函数存在多个瑕点和函数存在非正值的情况;
- 比较判别法、极限比较判别法和 p 判别法的总结;
- 常见函数在 ∞ 或 $-\infty$ 附近的变化;
- 常见函数在 0 附近的变化;
- 如何处理在非 0 有限值处的瑕点.

21.1 如 何 开 始

给定一个反常积分 $\int_a^b f(x)\mathrm{d}x$ (我们总是假设 f 是连续的或者有有限个不连续的点). 之所以称其为反常积分, 是因为被积函数 f 在区间 $[a,b]$ 上至少有一个瑕点. 瑕点经常出现在 f 的破裂点, 如有垂直渐近线的点处, 还出现在 ∞ 和 $-\infty$ 处. 例如, 积分

$$\int_{-\infty}^{\infty} \frac{1}{x^2-1}\mathrm{d}x$$

在 ∞ 和 $-\infty$ 处有瑕点 (只要包含它们, 就一定有瑕点), 同样 $x=1$ 和 $x=-1$ 处也有瑕点 (因为被积函数在这些点未定义).

如 20.1.2 节所述, 每次只关注一个瑕点是合理的. 同样地, 我们倾向于被积函数恒正, 至少 x 在瑕点附近时函数应该为正. 因此, 我们的第一个任务是适当地拆分积分, 第二个任务是处理 f 存在负值的情况.

21.1.1 拆分积分

下面是基本的对策：

(1) **确定**区间 $[a,b]$ 上的**所有瑕点**;

(2) **将积分拆分**成若干积分之和, 使得每个积分至多有一个瑕点, 并使这些瑕点作为相应积分的上限或下限;

(3) **分别讨论每个积分, 如果某一积分发散, 则整个积分发散**. 原反常积分收敛的唯一情形是, 每个积分都收敛.

如何将一个积分正确拆分呢? 如果只在 a 或 b 点有瑕点, 则什么都不用做. 考虑如下的经典例子, 积分

$$\int_0^\infty \frac{1}{\sqrt{x}+x^2}\mathrm{d}x$$

收敛还是发散? 被积函数在 $x=0$ 有垂直渐近线, ∞ 总是瑕点, 所以我们有端点 0 和 ∞ 两个瑕点. 这是两个瑕点, 而我们每个积分只处理一个瑕点, 所以可以在 0 和 ∞ 之间任选一个你喜欢的数字, 我选的是 5, 然后把积分拆分成两个:

$$\int_0^5 \frac{1}{\sqrt{x}+x^2}\mathrm{d}x \quad 和 \quad \int_5^\infty \frac{1}{\sqrt{x}+x^2}\mathrm{d}x.$$

这个例子将在 21.4.1 节完成. 现在, 注意这两个积分都只有一个瑕点, 并且瑕点在积分区间的左端点或右端点. 在哪个点对积分进行拆分是没有关系的, 这在 20.1.1 节已经详细讨论过了. 你可以用 20.1.2 节的例子进行验证, 那个反常积分需拆分成 5 个积分.

至于上面的第 (3) 步, 本章余下部分会探讨如何处理只在一个端点处有一个瑕点的积分. 为了让原积分收敛, 重要的是拆分后的每个积分必须都收敛. 所以若将一个反常积分拆分成 5 个积分, 发现有一个积分发散, 则不需要浪费时间考虑其他 4 个积分, 因为你已经知道整个积分是发散的了.

有一个重要的情形: 如果没有瑕点会怎样呢? 也就是说, 假设有积分 $\int_a^b f(x)\mathrm{d}x$, 它的积分区间 $[a,b]$ 是有界的 (故没有 ∞ 和 $-\infty$), 并且 f 在闭区间 $[a,b]$ 上有界, 则如 20.1 节所述, f 没有瑕点, 所以我们知道积分 $\int_a^b f(x)\mathrm{d}x$ 收敛. 总之, **如果没有瑕点, 则积分收敛!** 例如:

$$\int_0^{100} \frac{\ln(x+1)}{x^4+x^2+1}\mathrm{d}x$$

收敛, 因为被积函数在有界区间 $[0,100]$ 上有界, 也就是函数在该闭区间上没有瑕点. 对于形如上例的积分, 不要被其蒙蔽而用相关的判别法进行积分敛散性的判定.

21.1.2 如何处理负函数值

如果 $f(x)$ 在区间 $[a,b]$ 上的某些 x 处取负值, 你需要特别小心, 这经常出现在三角函数或对数函数中. 幸运的是, 你能够将问题化简为只有正被积函数的积分. 下面是处理负函数值的三种方法.

(1) 如果被积函数 $f(x)$ 在区间 $[a,b]$ 上既有正值又有负值, 应该考虑使用绝对收敛判别法. 如 20.6 节所述:

如果 $\int_a^b |f(x)|\mathrm{d}x$ 收敛, 那么 $\int_a^b f(x)\mathrm{d}x$ 也收敛.

这个判别法特别适用于讨论当积分区间不是有界的, 且包含三角函数的反常积分, 20.6 节的例子

$$\int_1^\infty \frac{\sin(x)}{x^2}\mathrm{d}x$$

就是这种类型的. 记住要从积分的绝对值开始考虑, 即

$$\int_1^\infty \frac{|\sin(x)|}{x^2}\mathrm{d}x;$$

不需要在分母上加绝对值, 因为分母恒正. 然后指出这个新的积分是收敛的 (详见 20.6 节), 并用绝对收敛判别法得出原积分也收敛. 不要忘记: 绝对收敛判别法只能帮你判断积分收敛, **不能用绝对收敛判别法判断积分发散!**

(2) 假设被积函数 $f(x)$ 在区间 $[a,b]$ 上恒负 (或为 0), 即在 $[a,b]$ 上 $f(x)\leqslant 0$. 则

$$\int_a^b f(x)\mathrm{d}x = -\int_a^b (-f(x))\mathrm{d}x$$

又会怎样? 现在 $-f(x)$ 非负, 所以可以用比较判别法或 p 判别法来看 $\int_a^b (-f(x))\mathrm{d}x$ 收敛还是发散. 当然, 如果该积分收敛, 则 $\int_a^b f(x)\mathrm{d}x$ 收敛. 类似地, 如果 $\int_a^b (-f(x))\mathrm{d}x$ 发散, 则 $\int_a^b f(x)\mathrm{d}x$ 发散. 例如

$$\int_0^{1/2} \frac{1}{x^2\ln(x)}\mathrm{d}x$$

显然在 $x=0$ 处有一个瑕点. 注意 $\ln(x)$ 在定义域 0 和 1 之间是负的, 所以最好写成

$$\int_0^{1/2} \frac{1}{x^2\ln(x)}\mathrm{d}x = -\int_0^{1/2} \frac{-1}{x^2\ln(x)}\mathrm{d}x.$$

事实上, 因为 $\ln(x)$ 是出现负值的部分, 所以可以将 $-\ln(x)$ 替换为 $|\ln(x)|$, 即

$$\int_0^{1/2} \frac{1}{x^2\ln(x)}\mathrm{d}x = -\int_0^{1/2} \frac{1}{x^2|\ln(x)|}\mathrm{d}x.$$

现在我们只需考虑

$$\int_0^{1/2} \frac{1}{x^2|\ln(x)|}\mathrm{d}x.$$

遗憾的是, 还要等到学习了 21.4.4 节才能最后知道这个积分是发散的, 因而原积分也发散. 注意绝对收敛判别法不能用于这个例子, 因为该判别法只能用于反常积分

收敛的判别.

(3) 如果上面两种情形都不适用, 可以用反常积分的正式定义试一下. 例如 20.6 节的

$$\int_0^\infty \cos(x)\mathrm{d}x.$$

到这里其实还没完, 还有一些特殊的反常积分收敛, 但不绝对收敛[1]. 这些类型的反常积分经常在实际的物理和工程应用中出现, 不过在本书的讨论范围之外. 现在我们来回顾一下积分判别法.

21.2 积分判别法总结

目前你可以使用的最有价值的工具是比较判别法、极限比较判别法和 p 判别法. 上一章我们从理论的角度讨论了这些判别法, 现在我们再次讨论它们. **在下面的所有判别法中, 被积函数 $f(x)$ 被假定为在积分区间上恒正.**

- **比较判别法的发散情形:** 若认为 $\int_a^b f(x)\mathrm{d}x$ 发散, 那就去找一个积分也发散的较小函数, 即找一个使得在区间 (a,b) 上有 $f(x) \geqslant g(x)$ 的非负函数 $g(x)$, 且 $\int_a^b g(x)\mathrm{d}x$ 发散. 则有
$$\int_a^b f(x)\mathrm{d}x \geqslant \int_a^b g(x)\mathrm{d}x = \infty,$$
因此 $\int_a^b f(x)\mathrm{d}x$ 发散.
- **比较判别法的收敛情形:** 若认为 $\int_a^b f(x)\mathrm{d}x$ 收敛, 那就去找一个积分也收敛的较大函数, 即找一个使得在区间 (a,b) 上有 $f(x) \leqslant g(x)$ 的函数 $g(x)$, 且 $\int_a^b g(x)\mathrm{d}x$ 收敛. 则有
$$\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x < \infty,$$
因此 $\int_a^b f(x)\mathrm{d}x$ 也收敛.

要小心, 别做无用功. 在 20.3 节中讨论过, 若搞反了上述不等式的方向就会做无用功. 若不等式的方向是反的, 就不能发挥比较判别法的作用.

极限比较判别法是比较判别法的替代形式. 使用该判别法的重点是, 能找到一个和被积函数在瑕点附近敛散性一致的函数. 在 20.4.1 节中, 我们有如下的定义:

$$\boxed{\text{当 } x \to a \text{ 时, } f(x) \sim g(x), \text{ 这等价于 } \lim_{x\to a}\frac{f(x)}{g(x)} = 1.}$$

若将 $x \to a$ 换成 $x \to \infty$ (或 $x \to -\infty$), 上述定义仍成立. 在任何情况下, 如

[1] 例如 $\int_1^\infty \sin(x)/x\mathrm{d}x$ 收敛, 但 $\int_1^\infty |\sin(x)|/x\mathrm{d}x$ 发散. 如果你能对它们中任一个的正确性加以说明, 你已经相当了不起了!

果被积函数 f 形式复杂, 而又能找到一个好的函数 g, 使得当 x 趋近于瑕点时有 $f(x) \sim g(x)$, 则你已经接近成功了! 这是因为根据极限比较判别法, g 与 f 敛散性一致. 更准确地, 下面是该判别法针对瑕点的有限和无穷两种情形的判别.

- **极限比较判别法中瑕点为无穷的情形**: 找一个在区间 $[a,\infty)$ 上没有瑕点、形式较简单的非负函数 g, 且有当 $x \to \infty$ 时, $f(x) \sim g(x)$, 则
  (1) 若 $\int_a^\infty g(x)\mathrm{d}x$ 收敛, 则 $\int_a^\infty f(x)\mathrm{d}x$ 收敛;
  (2) 若 $\int_a^\infty g(x)\mathrm{d}x$ 发散, 则 $\int_a^\infty f(x)\mathrm{d}x$ 发散.

当然, 将区间 $[a,\ \infty)$ 换为 $(-\infty,\ b]$ 也成立. 还有一种情形也成立, 即瑕点为积分区间左端点处的有限值 a.

- **极限比较判别法中瑕点为有限值的情形**: 找一个在区间 $(a,\ b]$ 上没有瑕点、形式较简单的非负函数 g, 且有当 $x \to a$ 时, $f(x) \sim g(x)$, 则
  (1) 若 $\int_a^b g(x)\mathrm{d}x$ 收敛, 则 $\int_a^b f(x)\mathrm{d}x$ 收敛;
  (2) 若 $\int_a^b g(x)\mathrm{d}x$ 发散, 则 $\int_a^b f(x)\mathrm{d}x$ 发散.

不用说, 针对唯一的瑕点在右端点 $x = b$, 且有当 $x \to b$(而不是 a) 时 $f(x) \sim g(x)$ 的情形, 结论相同.

因此, 需要我们找到一个合适的函数 g 来做比较. 通过选择 $g(x)$ 为 $1/x^p$ 的形式, 并选择合理的 p 值, 能够解决很多问题. 这类函数积分的敛散性可以准确地由 p 判别法描述.

- **p 判别法, $\int^\infty$ 的情形**: 对任意有限值 $a > 0$, 积分

$$\boxed{\int_a^\infty \frac{1}{x^p}\mathrm{d}x,\ \text{当 } p > 1 \text{ 时收敛, 当 } p \leqslant 1 \text{ 时发散;}}$$

- **p 判别法, $\int_0$ 的情形**: 对任意有限值 $a > 0$, 积分

$$\boxed{\int_0^a \frac{1}{x^p}\mathrm{d}x,\ \text{当 } p < 1 \text{ 时收敛, 当 } p \geqslant 1 \text{ 时发散.}}$$

好好学习所有这些判别法, 它们都是你的朋友.

21.3 常见函数在 ∞ 和 $-\infty$ 附近的表现

现在该回答最重要的问题了: 如何选择用于比较的函数 g? 这取决于瑕点在 $\pm\infty$、0, 还是其他的有限值处, 我们将分别讨论. 在几乎所有要讨论的情形中, 我们要重述之前的极限和不等式, 应用这些原理来讨论反常积分. 现在, 我们讨论常见函数在 ∞ 和 $-\infty$ 附近的情形.

21.3.1 多项式和多项式型函数在 ∞ 和 $-\infty$ 附近的表现

自多项式被研究以来, 在 $x \to \infty$ 或 $x \to -\infty$ 时**最高次项**起决定作用. 更准确地说, 设 p 为多项式, 则

若 $p(x)$ 的最高次项是 ax^n, 则当 $x \to \infty$ 或 $x \to -\infty$ 时, 有 $p(x) \sim ax^n$.

例如, 我们有

$$\text{当 } x \to \infty \text{ 时,} \quad x^5 + 4x^4 + 1 \sim x^5.$$

不用这种说法, 也可以通过指出当 $x \to \infty$ 时, $x^5 + 4x^4 + 1$ 和 x^5 的商的极限为 1 来验证. 过程如下:

$$\lim_{x\to\infty} \frac{x^5 + 4x^4 + 1}{x^5} = \lim_{x\to\infty} \left(1 + \frac{4}{x} + \frac{1}{x^5}\right) = 1.$$

在 4.3 节, 我们讨论了上述原理.

若 p 是一个多项式型函数而不是多项式, 则有一个类似原理适用. (欲知有关多项式型函数更多的信息, 参见 4.4 节.) 例如, 为了解 $x \to \infty$ 时的 $3\sqrt{x} - 2\sqrt[3]{x} + 4$, 将它写为 $3x^{1/2} - 2x^{1/3} + 4$, 由于最高次幂为 $1/2$, 我们可以说当 $x \to \infty$ 时, $3\sqrt{x} - 2\sqrt[3]{x} + 4 \sim 3\sqrt{x}$. (当 $x \to -\infty$ 时不成立, 因为负数不能开平方!)

有时最高次幂不好确定. 例如, $\sqrt{x^4 + 8x^3 - 9} - x^2$ 看起来是一个有最高次幂 4 的关于 x 的多项式型函数, 不过要开平方, 就会使幂次下降为 2. 当将 x^2 项消掉后, 最高次幂就有些难以理解了. 在本节末, 我们将讨论如何处理这样的问题.



由于我们有许多新的渐近关系, 故可以用极限比较判别法分析很多反常积分. 例如, 考虑

$$\int_1^\infty \frac{1}{2 + 20\sqrt{x}}\mathrm{d}x \quad \text{和} \quad \int_0^\infty \frac{1}{x^5 + 4x^4 + 1}\mathrm{d}x.$$

在这两个积分中, ∞ 都是唯一的瑕点. 第一个积分的分母 $2 + 20\sqrt{x}$ 可以写为 $2 + 20x^{1/2}$, 这里 $1/2$ 是最高次幂, 因此当 $x \to \infty$ 时, 有 $2 + 20\sqrt{x} \sim 20x^{1/2}$, 则

$$\text{当 } x \to \infty \text{ 时,} \quad \frac{1}{2 + 20\sqrt{x}} \sim \frac{1}{20x^{1/2}}.$$

现在由 p 判别法可知积分

$$\int_1^\infty \frac{1}{20x^{1/2}}\mathrm{d}x$$

发散, 由极限比较判别法可知积分

$$\int_1^\infty \frac{1}{2 + 20\sqrt{x}}\mathrm{d}x$$

也发散. 对于第二个积分, 由于当 $x \to \infty$ 时, 有 $x^5 + 4x^4 + 1 \sim x^5$, 则对倒数也一样有

$$\text{当 } x \to \infty \text{ 时}, \quad \frac{1}{x^5 + 4x^4 + 1} \sim \frac{1}{x^5}.$$

这里需要当心! 我们希望讨论的积分与积分 $\int_0^\infty 1/x^5 \mathrm{d}x$ 表现一样, 但问题是这个积分在 $x = 0$ 还有一个瑕点. 事实上, 该积分只因 0 点的瑕点而发散, 这将导致整个结论错误. 为了避免这些错误, 我们需要将原积分分成两部分:

$$\int_0^1 \frac{1}{x^5 + 4x^4 + 1} \mathrm{d}x \quad \text{和} \quad \int_1^\infty \frac{1}{x^5 + 4x^4 + 1} \mathrm{d}x.$$

在这两个积分中, 第一个因没有瑕点而收敛. 对于第二个积分, 我们有

$$\text{当 } x \to \infty \text{ 时}, \quad \frac{1}{x^5 + 4x^4 + 1} \sim \frac{1}{x^5}.$$

由于 $\int_1^\infty 1/x^5 \mathrm{d}x$ 收敛, 则积分

$$\int_1^\infty \frac{1}{x^5 + 4x^4 + 1} \mathrm{d}x$$

也收敛.

两个积分都收敛, 所以原积分也收敛. 这种情况经常出现, 所以要小心, 记着确保将积分进行拆分. 基本上, 如果 "极限比较函数" g 有原函数没有的瑕点, 为了避免产生新的瑕点, 你需要将原函数进行拆分. 通常新的被积函数 $g(x)$ 具有形式 $1/x^p$, 所以当有瑕点 ∞ 时, 就像我们例子一样, 只需避免出现 $x = 0$.

我们来看另一个例子

$$\int_2^\infty \frac{3x^5 + 2x^2 + 9}{x^6 + 22x^4 + \sqrt{4x^{13} + 18x}} \mathrm{d}x.$$

这个问题有点复杂, 唯一的瑕点是 ∞. 被积函数的分子很容易处理: 当 $x \to \infty$, 有 $3x^5 + 2x^2 + 9 \sim 3x^5$. 对于分母, 首先注意到当 $x \to \infty$, $\sqrt{4x^{13} + 18x} \sim \sqrt{4x^{13}} = 2x^{13/2}$. 由于 $13/2$ 大于 6, $\sqrt{4x^{13} + 18x}$ 项在分母中起主要作用, 所以当 $x \to \infty$, 整个分母渐近等价于 $2x^{13/2}$. 综上所述, 我们有

$$\text{当} x \to \infty \text{时}, \quad \frac{3x^5 + 2x^2 + 9}{x^6 + 22x^4 + \sqrt{4x^{13} + 18x}} \sim \frac{3x^5}{2x^{13/2}} = \frac{3}{2} \frac{1}{x^{3/2}}$$

由 p 判别法可知积分

$$\frac{3}{2} \int_2^\infty \frac{1}{x^{3/2}} \mathrm{d}x$$

收敛, 所以由极限比较判别法可知原积分也收敛.

最后, 考虑

$$\int_9^\infty \frac{1}{\sqrt{x^4 + 8x^3 - 9} - x^2} \mathrm{d}x.$$

如上面的讨论, 由于 $\sqrt{x^4}$ 与 x^2 相消, 分母的最高次幂难以确定, 我们需将分子分母同时乘以分母的共轭表达式. (之前已多次用过此法, 更多例题见 4.2 节.) 我们有

$$\int_9^{\infty}\frac{1}{\sqrt{x^4+8x^3-9}-x^2}\mathrm{d}x=\int_9^{\infty}\frac{1}{\sqrt{x^4+8x^3-9}-x^2}\times\frac{\sqrt{x^4+8x^3-9}+x^2}{\sqrt{x^4+8x^3-9}+x^2}\mathrm{d}x;$$



可将其化简为

$$\int_9^{\infty}\frac{\sqrt{x^4+8x^3-9}+x^2}{8x^3-9}\mathrm{d}x.$$

分母很容易处理：当 $x\to\infty$ 时, $8x^3-9\sim 8x^3$. 分子呢？由于 $x^4+8x^3-9\sim x^4$, 有 $\sqrt{x^4+8x^3-9}\sim x^2$, 最后 $\sqrt{x^4+8x^3-9}+x^2\sim 2x^2$ (当 $x\to\infty$ 时). 最后的结论有点难以理解, 因为渐近等价不能相加或相减. 为了确定该说法的正确性, 我们需要指出 $\sqrt{x^4+8x^3-9}+x^2$ 和 $2x^2$ 的比值当 $x\to\infty$ 时趋于 1, 因为

$$\lim_{x\to\infty}\frac{\sqrt{x^4+8x^3-9}+x^2}{2x^2}=\lim_{x\to\infty}\frac{1}{2}\left(\frac{\sqrt{x^4+8x^3-9}}{x^2}+\frac{x^2}{x^2}\right).$$

将分母上的 x^2 拖入根式 (为 x^4) 并化简, 上述极限变为

$$\begin{aligned}\lim_{x\to\infty}\frac{1}{2}\left(\sqrt{\frac{x^4+8x^3-9}{x^4}}+1\right)&=\lim_{x\to\infty}\frac{1}{2}\left(\sqrt{1+\frac{8}{x}-\frac{9}{x^4}}+1\right)\\&=\frac{1}{2}(\sqrt{1+0-0}+1)=1.\end{aligned}$$

这就证明了当 $x\to\infty$ 时, $\sqrt{x^4+8x^3-9}+x^2\sim 2x^2$. 我们再回到原积分, 得到

$$\frac{1}{\sqrt{x^4+8x^3-9}-x^2}=\frac{\sqrt{x^4+8x^3-9}+x^2}{8x^3-9}\sim\frac{2x^2}{8x^3}=\frac{1}{4x},\ \text{当 } x\to\infty.$$

运用极限比较判别法, 由于 $\int_9^{\infty}1/(4x)\mathrm{d}x$ 发散, 因而原积分发散. 顺便说一下, 你能猜到原被积函数在 $x\to\infty$ 时渐近等价于 $1/4x$ 吗？这个不容易想到, 所以如果要用最高次幂起决定作用的结论, 要保证有且仅有一个最高次幂.

21.3.2 三角函数在 ∞ 和 $-\infty$ 附近的表现

或许我们能知道的有用结论仅仅是, 对任意实数 A 有

$$\boxed{|\sin(A)|\leqslant 1}\quad \text{和}\quad \boxed{|\cos(A)|\leqslant 1.}$$



虽然给出的信息不多, 但总比没有好. (其他的三角函数有太多的垂直渐近线, 所以它们不满足类似的不等式.) 上述不等式有两个主要的应用, 一是可以在很多情况下使用比较判别法. 例如, 积分

$$\int_5^{\infty}\frac{|\sin(x^4)|}{\sqrt{x}+x^2}\mathrm{d}x$$

收敛还是发散呢？我们从 $|\sin(x^4)|\leqslant 1$ 开始. 注意, 将 A 的正弦值换成 x^4 的正弦值是没关系的, 因为任何数的正弦 (或余弦) 的绝对值都不超过 1. 因此, 我们有

$$\int_5^\infty \frac{|\sin(x^4)|}{\sqrt{x}+x^2}\mathrm{d}x \leqslant \int_5^\infty \frac{1}{\sqrt{x}+x^2}\mathrm{d}x.$$

太棒了, 我们去除了表达式中的所有三角函数, 右边积分的唯一瑕点出现在 ∞ 处. 由于对于大数 x, 最高次幂起主要作用, 所以我们有当 $x \to \infty$, $\sqrt{x}+x^2 \sim x^2$. 现在取倒数可得

$$\text{当 } x \to \infty \text{ 时,} \quad \frac{1}{\sqrt{x}+x^2} \sim \frac{1}{x^2}.$$

由 p 判别法, 我们知道 $\int_5^\infty 1/x^2\mathrm{d}x$ 收敛, 极限比较判别法告诉我们

$$\int_5^\infty \frac{1}{\sqrt{x}+x^2}\mathrm{d}x$$

也收敛. 最后, 我们有

$$\int_5^\infty \frac{|\sin(x^4)|}{\sqrt{x}+x^2}\mathrm{d}x \leqslant \int_5^\infty \frac{1}{\sqrt{x}+x^2}\mathrm{d}x < \infty,$$

所以由比较判别法知原积分收敛.

$|\sin(A)| \leqslant 1$ 和 $|\cos(A)| \leqslant 1$ 的另一个漂亮的应用是, 相对于 x 的任何正数次幂, 任何数的正弦或余弦值都可忽略, 至少在 $x \to \infty$ 或 $x \to -\infty$ 时是这样的. 例如

$$\text{当 } x \to \infty \text{ 时,} \quad 2x^3 - 3x^{0.1} + \sin(100x^{200}) \sim 2x^3.$$

为什么? 因为当 x 是大数时, 正弦项与 $2x^3$ 相比相当地小. 更准确地, 我们有

$$\lim_{x\to\infty} \frac{2x^3 - 3x^{0.1} + \sin(100x^{200})}{2x^3} = \lim_{x\to\infty}\left(1 - \frac{3}{2x^{2.9}} + \frac{\sin(100x^{200})}{2x^3}\right).$$

项 $3/2x^{2.9}$ 当 $x \to \infty$ 时趋于 0. 关键点是, 可以用三明治定理得出

$$\lim_{x\to\infty} \frac{\sin(100x^{200})}{2x^3} = 0.$$



具体过程留给你来完成, 因为我们在 7.1.3 节讨论过类似的例子. 不管怎样, 我们已经得出

$$\lim_{x\to\infty} \frac{2x^3 - 3x^{0.1} + \sin(100x^{200})}{2x^3} = 1.$$

毕竟这就证得

$$\text{当 } x \to \infty \text{ 时,} \quad 2x^3 - 3x^{0.1} + \sin(100x^{200}) \sim 2x^3.$$



这个结论对于了解积分

$$\int_8^\infty \frac{1}{2x^3 - 3x^{0.1} + \sin(100x^{200})}\mathrm{d}x$$

收敛与否是很有用的. 由极限比较判别法和上面的渐近等价关系可知, 该积分与 $\int_8^\infty 1/2x^3\mathrm{d}x$ 敛散性一致. 因为据 p 判别法, 最后一个积分收敛, 则原积分也收敛.

21.3.3 指数在 ∞ 和 $-\infty$ 附近的表现

这是一个非常有用的原理：**指数比多项式增长得快**. 我们在 9.4.4 节最先给出这个结论, 当时用下面的形式来表示这个原理：

$$\lim_{x\to\infty}\frac{x^n}{\mathrm{e}^x}=0,$$

其中 n 是任意正数, 甚至很大的数. 现在考虑函数 $f(x)=x^n/\mathrm{e}^x$, 我们可知 $f(0)=0$, 且由上面的极限有当 $x\to\infty$ 时 $f(x)\to 0$. 那么当 $x\geqslant 0$ 时, $f(x)$ 能有多大呢? 函数从 0 开始, 中间没有垂直渐近线, 然后又折返下来, 在 $y=0$ 处有水平渐近线, 所以 $y=f(x)$ 的图像必然有最大高度, 我们定义为 C. 意思是对所有的 $x\geqslant 0$, $f(x)=x^n/\mathrm{e}^x\leqslant C$. (注意对不同的 n 有不同的 C 与之对应, 但这无关紧要.) 现在, 将 $1/\mathrm{e}^x$ 写为 e^{-x}, 并两边同时除以 x^n, 得到一个有用的不等式

$$\boxed{\text{对所有的 } x>0,\ \mathrm{e}^{-x}\leqslant\frac{C}{x^n}.}$$

如 9.4.4 节所述, 如果将 e^{-x} 换成 $\mathrm{e}^{-p(x)}$ 也是对的, 这里 $p(x)$ 是当 $x\to\infty$ 时趋于无穷的任何一个多项式型表达式, 底数 e 也可以换成其他大于 1 的数. 例如, 若将 e^{-x} 换为 $2^{-5x^5+\sqrt{x^3+3}}$, 上述不等式也成立. 这里重点是你可以任意选择 n, 但要注意使它足够地大. 例如, 考虑



$$\int_1^\infty x^3\mathrm{e}^{-x}\mathrm{d}x.$$

好消息是, 被积函数是正的且只有 ∞ 一个瑕点; 坏消息是, 因子 x^3 在 $x\to\infty$ 时增长很快. 然而因子 e^{-x} 减小 (到 0) 非常快, 其速度要远远快于 x^3 的增长速度. 为了证明这个, 我们将关注

$$\mathrm{e}^{-x}\leqslant\frac{C}{x^5}.$$

这正好是方框中的不等式, 只是将 n 选为 5, 为什么选 5 呢? 因为它能起作用：

$$\int_1^\infty x^3\mathrm{e}^{-x}\mathrm{d}x\leqslant\int_1^\infty x^3\frac{C}{x^5}\mathrm{d}x=C\int_1^\infty\frac{1}{x^2}\mathrm{d}x<\infty.$$

我们已经用 p 判别法得到 $C\int_1^\infty 1/x^2\mathrm{d}x$ 收敛. 由比较判别法可知, 原积分也收敛. 我是怎么知道用 x^5 的呢? 如果换成 $\mathrm{e}^{-x}\leqslant C/x^4$ 会发生什么? 它将不会起作用：

$$\int_1^\infty x^3\mathrm{e}^{-x}\mathrm{d}x\leqslant\int_1^\infty x^3\frac{C}{x^4}\mathrm{d}x=C\int_1^\infty\frac{1}{x}\mathrm{d}x=\infty.$$



我们完全白费功夫了, 因为除了能够说明原积分有穷或无穷外, 其他什么都说明不了. 另一方面, 如果我们之前用了 $x^{4.0001}$ 就会有用, 为什么? 你只要保证所选的指数为比 4 大的任意数, 该论证就能有作用. 实际上, 最好选择要消除的幂加 2 的数.


这里我们想消去 x^3, 所以用 $\mathrm{e}^{-x} \leqslant C/x^5$.


重要的一点是: 若说当 $x \to \infty$, $x^3\mathrm{e}^{-x} \sim \mathrm{e}^{-x}$ 就大错特错了. 它是不正确的. 如果是正确的, 就可以消掉正项 e^{-x} 而得到结论当 $x \to \infty$, $x^3 \sim 1$. 这才是瞎说. 所以, 你应该对前一个例子采用比较判别法而不是极限比较判别法.

现在来看积分

$$\int_{10}^{\infty}(x^{1000}+x^2+\sin(x))\mathrm{e}^{-x^2+6}\mathrm{d}x.$$

我们要做一点点的工作. 被积函数因有 $\sin(x)$ 项而看起来在正值和负值之间振荡, 不过这不是事实, 因为 $\sin(x)$ 的大小不足以影响正数 $x^{1000}+x^2(x \geqslant 10)$ 的符号. 不管怎样, 第一个观察的结果是当 $x \to \infty$, $x^{1000}+x^2+\sin(x) \sim x^{1000}$, 因为 x^2 和 $\sin(x)$ 项的作用被 x^{1000} 抵掉了. (若想知道如何给出更专业的解释, 见前一节.) 故我们乘以 e^{-x^2+6} 得

$$\text{当 } x \to \infty \text{ 时,} \quad (x^{1000}+x^2+\sin(x))\mathrm{e}^{-x^2+6} \sim x^{1000}\mathrm{e}^{-x^2+6}.$$

利用极限比较判别法, 我们只需知道积分

$$\int_{10}^{\infty} x^{1000}\mathrm{e}^{-x^2+6}\mathrm{d}x$$

收敛还是发散. 原积分也需要做相同讨论. 现在需要小心了, 因为指数项 e^{-x^2+6} 没有合适的渐近等价对象, 这里我们需要采用基本的比较方法. 你知道, x^{1000} 确实增加了, 但 e^{-x^2+6} 的的确确在减小. 我们用

$$\mathrm{e}^{-x^2+6} \leqslant \frac{C}{x^{1002}}$$

(看, 1002 比 1000 大 2) 来得到

$$x^{1000}\mathrm{e}^{-x^2+6} \leqslant x^{1000} \times \frac{C}{x^{1002}} = \frac{C}{x^2}.$$

故用比较判别法有

$$\int_{10}^{\infty} x^{1000}\mathrm{e}^{-x^2+6}\mathrm{d}x \leqslant C\int_{10}^{\infty}\frac{1}{x^2}\mathrm{d}x < \infty$$

(其中最后那个积分由 p 判别法得知收敛). 整理一下思路. 我们知道积分

$$\int_{10}^{\infty} x^{1000}\mathrm{e}^{-x^2+6}\mathrm{d}x$$

收敛, 因此由极限比较判别法知

$$\int_{10}^{\infty}(x^{1000}+x^2+\sin(x))\mathrm{e}^{-x^2+6}\mathrm{d}x$$

也收敛.


e^x 在 $-\infty$ 附近是什么表现呢? 这与讨论 e^{-x} 在 ∞ 附近的表现是一回事. 例如, 考虑

$$\int_{-\infty}^{-4} x^{1000}\mathrm{e}^x \mathrm{d}x,$$

首先做变量代换 $t=-x$, 由 $\mathrm{d}t=-\mathrm{d}x$, 我们有

$$\int_{-\infty}^{-4} x^{1000}\mathrm{e}^x \mathrm{d}x = -\int_{\infty}^{4} (-t)^{1000}\mathrm{e}^{-t}\mathrm{d}t = \int_{4}^{\infty} t^{1000}\mathrm{e}^{-t}\mathrm{d}t.$$



这里我们用由 $\mathrm{d}t$ 产生的负号将积分的上下限进行了调换. 最后这个积分的敛散判别留给你们自己完成.



这里有个怪题：积分

$$\int_{4}^{\infty} x^{1000}\mathrm{e}^x \mathrm{d}x$$

收敛还是发散呢? 被积函数的两个因子在 $x\to\infty$ 时都无限增大, 所以它当然发散! 更准确地说, 当 $x\geqslant 4$, 显然 $x^{1000}\mathrm{e}^x\geqslant 1$ (事实上, 不等式右边的 1 是保守的选择). 则我们有

$$\int_{4}^{\infty} x^{1000}\mathrm{e}^x \mathrm{d}x \geqslant \int_{4}^{\infty} 1\mathrm{d}x = \infty.$$

一定要保证右边的积分发散. (这是不证自明的, 不过你可以用正式定义或 $p=0$ 时的 p 判别法加以证明). 总之, 比较判别法推出了原积分发散.



现在再来考虑指数加上多项式后会发生什么. 如你所希望的, 如果指数变得很大, 与多项式相比, 它的变化起决定作用. 例如, 分析

$$\int_{9}^{\infty} \frac{x^{10}}{\mathrm{e}^x-5x^{20}}\mathrm{d}x,$$

先看分母 e^x-5x^{20}, e^x 项与 $5x^{20}$ 项相比起决定作用, 我们应该有 $\mathrm{e}^x-5x^{20}\sim\mathrm{e}^x$ (当 $x\to\infty$). 我们可以通过讨论如下商式的极限来加以证明:

$$\lim_{x\to\infty}\frac{\mathrm{e}^x-5x^{20}}{\mathrm{e}^x} = \lim_{x\to\infty}\left(1-\frac{5x^{20}}{\mathrm{e}^x}\right) = 1-0 = 1.$$

(这里用了本节一开始的那个极限.) 总之, 由 $\mathrm{e}^x-5x^{20}\sim\mathrm{e}^x$(当 $x\to\infty$), 我们有

$$\text{当 } x\to\infty \text{ 时}, \quad \frac{x^{10}}{\mathrm{e}^x-5x^{20}} \sim \frac{x^{10}}{\mathrm{e}^x};$$

所以, 可以讨论

$$\int_{9}^{\infty} \frac{x^{10}}{\mathrm{e}^x}\mathrm{d}x = \int_{9}^{\infty} \mathrm{e}^{-x}x^{10}\mathrm{d}x,$$



来代替原积分. 接下来可用不等式 $\mathrm{e}^{-x}\leqslant C/x^{12}$ 和比较判别法证明该积分收敛, 这部分由你自己完成. 所以, 由极限比较判别法知原积分收敛.



最后, 考虑积分

$$\int_{18}^{\infty} \frac{x^2}{7^x-4^x}\mathrm{d}x.$$

我们最好知道分母 7^x-4^x 到底是什么. 这里 7^x 和 4^x 都是指数. 具有最大底数的起决定作用; 也就是说, 当 $x\to\infty$, $7^x-4^x\sim 7^x$. 为了说明原因, 来看它们比式的极限:

$$\lim_{x\to\infty}\frac{7^x-4^x}{7^x}=\lim_{x\to\infty}\left(1-\frac{4^x}{7^x}\right)=\lim_{x\to\infty}\left(1-\left(\frac{4}{7}\right)^x\right).$$

在 9.4.4 节有

$$\text{若}0\leqslant r<1,\quad \lim_{x\to\infty}r^x=0.$$

这是证明当 $x\to\infty$, $(4/7)^x\to 0$ 所需要的, 只需将 r 换成 4/7. 所以我们有

$$\lim_{x\to\infty}\frac{7^x-4^x}{7^x}=\lim_{x\to\infty}\left(1-\left(\frac{4}{7}\right)^x\right)=1-0=1.$$

这就证明了当 $x\to\infty$, $7^x-4^x\sim 7^x$. 所以, 我们也为原被积函数找到了一个渐近等价关系:



$$\text{当}x\to\infty,\quad \frac{x^2}{7^x-4^x}\sim\frac{x^2}{7^x}.$$

现在尝试用不等式 $7^{-x}\leqslant C/x^4$ 来证明

$$\int_{18}^{\infty}\frac{x^2}{7^x}\mathrm{d}x=\int_{18}^{\infty}7^{-x}x^2\mathrm{d}x$$

收敛, 故而由极限比较判别法可知原积分也收敛.

21.3.4　对数在 ∞ 附近的表现

首先注意我们不考虑对数在 $-\infty$ 附近的情形, 因为负数不能取对数, 所以讨论当 $x\to-\infty$ 时的 $\ln(x)$ 是没有意义的.

另一方面, **对数在∞处增长得很慢**. 事实上, 对数比 x 的任何正数次幂增长都慢. 用符号来表示就是, 若 $\alpha>0$ 是你选择的某个正数, 则不管它有多小, 都有

$$\lim_{x\to\infty}\frac{\ln(x)}{x^\alpha}=0.$$

在 9.4.5 节, 我们详细讨论了这个原理. 由 21.3.3 节的一个类似论述可得, 必有一个常数 C 使得

$$\boxed{\text{对所有 } x>1,\ \ln(x)\leqslant Cx^\alpha.}$$



上述结论对任何底数大于 1 的对数或最高次项系数为正的多项式的对数都成立.

例如, 考虑

$$\int_2^{\infty}\frac{\ln(x)}{x^{1.001}}\mathrm{d}x.$$

若没有 $\ln(x)$, 则由 p 判别法可知该积分收敛. 由于 $\ln(x)$ 增长得很慢, 它基本不会有什么影响. 虽然这是个正确概念, 但并不精确. 为了证明这个说法, 我们要用到

$\ln(x) \leqslant Cx^{\alpha}$, 其中 α 小到 x^{α} 不会影响 1.001 大于 1 这样一个性质. 例如, 如果采用 $\ln(x) \leqslant Cx^{0.5}$, 由 p 判别法可得

$$\int_2^{\infty} \frac{\ln(x)}{x^{1.001}}\mathrm{d}x \leqslant \int_2^{\infty} \frac{Cx^{0.5}}{x^{1.001}}\mathrm{d}x = C\int_2^{\infty} \frac{1}{x^{0.501}}\mathrm{d}x = \infty.$$

又白费力气了. 我们讨论的积分小于等于 ∞, 没有任何意义. 那我们就更精细一点, 采用 $\ln(x) \leqslant Cx^{0.0005}$. 0.0005 是一个很小的数, 小到被 1.001 减后的差仍大于 1. 我们看一下这个不等式的应用结果:

$$\int_2^{\infty} \frac{\ln(x)}{x^{1.001}}\mathrm{d}x \leqslant \int_2^{\infty} \frac{Cx^{0.0005}}{x^{1.001}}\mathrm{d}x = C\int_2^{\infty} \frac{1}{x^{1.0005}}\mathrm{d}x < \infty.$$

上面右边的积分根据 p 判别法可知是收敛的, 因为 1.0005 大于 1. 由比较判别法可知, 左边的积分也收敛. 你看到有多精细了吗? 这个方法与 21.3.3 节处理指数的方法类似.

提醒一下, 对数增长缓慢的原理并不是对每个含有对数的反常积分都有效. 考虑下面 6 个反常积分:

$$\int_2^{\infty} \frac{\ln(x)}{x^{1.001}}\mathrm{d}x, \quad \int_2^{\infty} \frac{1}{x^{1.001}\ln(x)}\mathrm{d}x, \quad \int_2^{\infty} \frac{\ln(x)}{x}\mathrm{d}x,$$

$$\int_2^{\infty} \frac{1}{x\ln(x)}\mathrm{d}x, \quad \int_{3/2}^{\infty} \frac{\ln(x)}{x^{0.999}}\mathrm{d}x, \quad \int_2^{\infty} \frac{1}{x^{0.999}\ln(x)}\mathrm{d}x.$$

我们刚讨论过第一个, 发现它是收敛的. 看第二个例子:

$$\int_2^{\infty} \frac{1}{x^{1.001}\ln(x)}\mathrm{d}x.$$

这里若没有因子 $\ln(x)$, 积分仍收敛, 但这个因子在分母上其实是有帮助的. 也就是说, 当 $\ln(x)$ 在分母上时, 分母变得比原来更大了, 使得被积函数变小了, 这有利于积分收敛. 如何更有效地把这些写下来呢? 随着 x 的增大, $\ln(x)$**有下界**. 这时, 积分区间是 $[2, \infty)$, 那么 $\ln(x)$ 在这个区间上能有多小呢? 由于 $\ln(x)$ 是关于 x 的增函数, 所以当 $x = 2$ 时, $\ln(x)$ 在该区间上有最小值. 所以, 我们仅需要写出当 $x \geqslant 2$ 时 $\ln(x) \geqslant \ln(2)$. 这有什么帮助吗? 两边取倒数, 发现当 $x \geqslant 2$ 有

$$\frac{1}{\ln(x)} \leqslant \frac{1}{\ln(2)},$$

然后两边同时除以 $x^{1.001}$ 后, 左边为被积函数:

$$\frac{1}{x^{1.001}\ln(x)} \leqslant \frac{1}{x^{1.001}\ln(2)}.$$

现在可用比较判别法, 因为

$$\int_2^{\infty} \frac{1}{x^{1.001}\ln(x)}\mathrm{d}x \leqslant \int_2^{\infty} \frac{1}{x^{1.001}\ln(2)}\mathrm{d}x = \frac{1}{\ln(2)}\int_2^{\infty} \frac{1}{x^{1.001}}\mathrm{d}x < \infty.$$



要知道 $\ln(2)$ 是一个常数, 因此可被提到积分符号前面, 由 p 判别法可知原积分收敛, 因为 1.001 比 1 大. 所以上面 6 个积分中的第二个积分收敛. 顺便说一下, 确定值 $\ln(2)$ 是无关紧要的, 我们可以将 $\ln(2)$ 换成任何常数 C, 证明仍然成立.

那么第三个积分呢? 看

$$\int_2^\infty \frac{\ln(x)}{x}\mathrm{d}x.$$

如果把分子中的 $\ln(x)$ 拿掉会怎样? 我们知道 $\int_2^\infty 1/x\mathrm{d}x$ 发散, 把 $\ln(x)$ 放回去, 情况只会变得更糟, 所以上述积分应该发散. 为了加以证明, 我们使用不等式 $\ln(x) \geqslant \ln(2)$, 此时 $x \geqslant 2$(或者可将 $\ln(2)$ 换成任何常数 $C > 0$). 可得

$$\int_2^\infty \frac{\ln(x)}{x}\mathrm{d}x \geqslant \int_2^\infty \frac{\ln(2)}{x}\mathrm{d}x = \ln(2)\int_2^\infty \frac{1}{x}\mathrm{d}x = \infty.$$



由比较判别法可知积分发散.

对第四个积分

$$\int_2^\infty \frac{1}{x\ln(x)}\mathrm{d}x,$$

现在需要做一些完全不同的事情了. 可以看到, 这个积分的任何部分都达到了完美的均衡. 如果没有因子 $\ln(x)$ 则积分发散. 由于 $\ln(x)$ 在分母中, 又会给积分以收敛的机会. 它的作用足够使积分收敛吗? 我们想利用 $\ln(x) \leqslant Cx^\alpha$, 但无论选择多么小的 α 都找不到一个有效的比较. (试一下就知道了!) 在此我们考虑用变量代换. 令 $t = \ln(x)$, 则 $\mathrm{d}t = 1/x\mathrm{d}x$. 当 $x = 2$, 我们有 $t = \ln(2)$, 且当 $x \to \infty$ 时有 $t \to \infty$, 所以



$$\int_2^\infty \frac{1}{x\ln(x)}\mathrm{d}x = \int_{\ln(2)}^\infty \frac{\mathrm{d}t}{t} = \infty,$$

其中后面的积分由 p 判别法可知发散, 则原积分也发散. 另一方面, 我们把上述积分的上限由 ∞ 换为 $\mathrm{e}^{\mathrm{e}^8}$:

$$\int_2^{\mathrm{e}^{\mathrm{e}^8}} \frac{1}{x\ln(x)}\mathrm{d}x.$$

数 $\mathrm{e}^{\mathrm{e}^8}$ 其实很大, 从我的电脑上获知它的值接近于 4×10^{1294}, 意味着 4 后面跟着 1294 个 0. 这是一个难以置信的大数, 相对于我们人脑的有限理解能力来说, 这个数就是无穷了. 因为当积分上限为 ∞ 时积分发散, 所以你可能认为上面积分的值是相当大的. 我们来计算它, 并令 $t = \ln(x)$, 可得

$$\int_2^{\mathrm{e}^{\mathrm{e}^8}} \frac{1}{x\ln(x)}\mathrm{d}x = \int_{\ln(2)}^{\mathrm{e}^8} \frac{1}{t}\mathrm{d}t = \ln(t)\Big|_{\ln(2)}^{\mathrm{e}^8} = \ln(\mathrm{e}^8) - \ln(\ln(2)) = 8 - \ln(\ln(2)).$$

这里我们用到了一点, 即当 $x = \mathrm{e}^{\mathrm{e}^8}$ 时有 $t = \ln(\mathrm{e}^{\mathrm{e}^8}) = \mathrm{e}^8$. 总之, 最终的结果比 8 小

一点点, 一点都不大. 这会使你认为反常积分

$$\int_2^\infty \frac{1}{x\ln(x)}\mathrm{d}x$$

收敛, 但如我们刚刚所见, 它是发散的, 只不过发散的速度非常缓慢.

现在考虑

$$\int_2^\infty \frac{1}{x(\ln(x))^{1.1}}\mathrm{d}x.$$

如果还用代换 $t=\ln(x)$, 可得

$$\int_2^\infty \frac{1}{x(\ln(x))^{1.1}}\mathrm{d}x = \int_{\ln(2)}^\infty \frac{\mathrm{d}t}{t^{1.1}} < \infty,$$

其中后一个积分由 p 判别法可知收敛, 所以新积分也收敛. 只需要让分母上的 $\ln(x)$ 的幂增加一点点, 即 $(\ln(x))^{0.1}$ 就足够让积分收敛了. 这一点增加真是重大的改变啊!

我们仍有两个积分需要考虑, 第一个是

$$\int_{3/2}^\infty \frac{\ln(x)}{x^{0.999}}\mathrm{d}x.$$

这个积分与第三个积分类似. 如果分子上没有因子 $\ln(x)$, 积分发散; 加上 $\ln(x)$ 只会使积分更加发散. 我们不能说对积分区间里的所有 x 都有 $\ln(x)\geqslant\ln(2)$, 因为现在的积分区间是 $[3/2,\ \infty)$. 不用管它, 只要换成 $\ln(x)\geqslant\ln(3/2)$ 就行了:

$$\int_{3/2}^\infty \frac{\ln(x)}{x^{0.999}}\mathrm{d}x \geqslant \int_{3/2}^\infty \frac{\ln(3/2)}{x^{0.999}}\mathrm{d}x = \ln(3/2)\int_{3/2}^\infty \frac{1}{x^{0.999}}\mathrm{d}x = \infty.$$

由 p 判别法可知最后一个积分发散. 又根据比较判别法, 原积分发散. (同样, 可以将 $\ln(3/2)$ 换成任意大于 0 的数 C.)

最后, 我们考虑本节的最后一个积分:

$$\int_2^\infty \frac{1}{x^{0.999}\ln(x)}\mathrm{d}x.$$

解这个题目的一个方法是直接利用比较判别法, 将其与第四个反常积分比较. 特别地, 当 $x\geqslant 2$ 时 $x^{0.999}<x$. 我们两边取倒数, 改变不等式的方向, 有

$$\int_2^\infty \frac{1}{x^{0.999}\ln(x)}\mathrm{d}x > \int_2^\infty \frac{1}{x\ln(x)}\mathrm{d}x.$$

我们已经知道上面最后一个积分是发散的, 所以由比较判别法可知原积分也发散. 还有一个更直接的方法. 观察原积分

$$\int_2^\infty \frac{1}{x^{0.999}\ln(x)}\mathrm{d}x,$$

如果把因子 $\ln(x)$ 拿走会发生什么呢? 根据 p 判别法可知它会发散. 把因子 $\ln(x)$ 放进分母会使积分有收敛的趋向, 但不是很明显. 事实上, 的确不足以使积分收敛. 你可以运用对数增长缓慢的原理: $\ln(x) \leqslant Cx^{0.0005}$, 两边取倒数, 我们有

$$\frac{1}{\ln(x)} \geqslant \frac{1}{C} \times \frac{1}{x^{0.0005}}.$$

不等式两边同时除以 $x^{0.999}$, 可得

$$\frac{1}{x^{0.999}\ln(x)} \geqslant \frac{1}{C} \times \frac{1}{x^{0.999}x^{0.0005}} = \frac{1}{C} \times \frac{1}{x^{0.9995}}.$$

最后可得

$$\int_2^{\infty} \frac{1}{x^{0.999}\ln(x)}\mathrm{d}x \geqslant \frac{1}{C}\int_2^{\infty} \frac{1}{x^{0.9995}}\mathrm{d}x = \infty,$$

最后一个积分由 p 判别法可知发散, 故原积分也发散. 注意我们仍选择足够小的幂次 0.0005, 其实我们还可以用任何小的正数, 只要当你把它加到 0.999 上不会得到大于等于 1 的数. 否则的话, 你又要白费力气了.

21.4　常见函数在 0 附近的表现

目前我们已经知道了, 多项式、三角函数、指数、对数在 ∞ 附近的表现. 再来看一下它们在 0 附近的表现.

21.4.1　多项式和多项式型函数在 0 附近的表现

对多项式, **最低次幂在 $x \to 0$ 时起决定作用**. 这与 $x \to \infty$ 时的情况正好相反. 更准确地, 假设 p 是多项式, 则有

若 $p(x)$ 的最低次项是 bx^m, 则当 $x \to 0$, $p(x) \sim bx^m$.

例如, 当 $x \to 0$, $5x^4 - x^3 + 2x^2 \sim 2x^2$. 我们通过证明它们之比的极限为 1 来说明:

$$\lim_{x\to 0} \frac{5x^4 - x^3 + 2x^2}{2x^2} = \lim_{x\to 0}\left(\frac{5x^2}{2} - \frac{x}{2} + 1\right) = 0 - 0 + 1 = 1.$$

对于多项式型函数, 并不是总能那么容易找到最低次项, 不过该原理仍适用. 例如, 当 $x \to 0^+$, $x^2 + \sqrt{x} \sim \sqrt{x}$, 因为 $\sqrt{x} = x^{1/2}$ 且 $1/2$ 小于 2. (这里 $x \to 0^+$, 因为不能对负数开平方.) 该原理甚至对常数也适用, 常数其实是 x^0 的倍数, 而 x^0 是次数很低的项. 如, 当 $x \to 0$, $2x^{1/3} + 4 \sim 4$, 因为 $4x^0$ 的指数低于 $2x^{1/3}$ 的指数.



我们来看一些关于反常积分的例子. 考虑

$$\int_0^5 \frac{1}{x^2 + \sqrt{x}}\mathrm{d}x.$$

唯一的瑕点是 $x = 0$, 现在我们知道

$$\text{当 } x \to 0^+ \text{ 时,}\quad \frac{1}{x^2 + \sqrt{x}} \sim \frac{1}{\sqrt{x}},$$

因为 $\int_0^5 1/\sqrt{x}\mathrm{d}x$ 收敛 (p 判别法), 则

$$\int_0^5 \frac{1}{x^2+\sqrt{x}}\mathrm{d}x$$

也收敛 (极限比较判别法). 因此积分收敛, 这主要是因为 $\sqrt{x}$ 项. 如果没有它, 被积函数为 $1/x^2$, 则积分在区间 [0, 5] 上发散, 所以 $\sqrt{x}$ 项保全了积分的收敛性. 不过等一下, 关于这一点, 我希望你回到 21.3.2 节看一下积分

$$\int_5^\infty \frac{1}{x^2+\sqrt{x}}\mathrm{d}x$$

是怎样收敛的. 后一个积分收敛的决定项是 x^2, 而不是 $\sqrt{x}$. 若没有 x^2, 后一个积分将会发散. 故我们在 21.1.1 节开始部分看到的积分

$$\int_0^\infty \frac{1}{x^2+\sqrt{x}}\mathrm{d}x$$

收敛, 因为以下两个积分

$$\int_0^5 \frac{1}{x^2+\sqrt{x}}\mathrm{d}x \quad 和 \quad \int_5^\infty \frac{1}{x^2+\sqrt{x}}\mathrm{d}x.$$

收敛. 瑕点 0 由于 $\sqrt{x}$ 项的存在没问题, 瑕点 ∞ 由于 x^2 项的存在也没问题, 非常不错, 是吧？



积分

$$\int_0^1 \frac{x+3}{x+x^5}\mathrm{d}x?$$

瑕点还是 $x=0$, 现在当 $x\to 0$, $x+3\sim 3$ 且 $x+x^5\sim x$, 故

$$当\ x\to 0\ 时, \quad \frac{x+3}{x+x^5}\sim\frac{3}{x},$$

反常积分 $\displaystyle\int_0^1 3/x\mathrm{d}x$ 由 p 判别法可知发散, 根据极限比较判别法可知原积分

$$\int_0^1 \frac{x+3}{x+x^5}\mathrm{d}x$$

也发散.

21.4.2 三角函数在 0 附近的表现

这些是很有用的结论：

$$\boxed{当\ x\to 0,\ \sin(x)\sim x,\ \tan(x)\sim x\ 且\ \cos(x)\sim 1.}$$

这些只是我们在第 7 章讨论过的极限的另一种描述：

$$\lim_{x\to 0}\frac{\sin(x)}{x}=1, \quad \lim_{x\to 0}\frac{\tan(x)}{x}=1, \quad \lim_{x\to 0}\cos(x)=1.$$

(若不明白余弦的极限, 将 $\cos(x)$ 写成 $\cos(x)/1$ 就可得当 $x \to 0$, $\cos(x) \sim 1$.) 注意：这些渐近等价关系的积和商成立, 而和与差不成立. 例如, 不能说当 $x \to 0$, $\sin(x) - x \sim 0$. 更深入的讨论见 20.4.1 节末.

我们来看一些例子. 考虑

$$\int_0^1 \frac{1}{\tan(x)}\mathrm{d}x \quad 和 \quad \int_0^1 \frac{1}{\sqrt{\tan(x)}}\mathrm{d}x.$$

这两个积分看上去很相似, 外表很有迷惑性. 我们对两个积分都采用 $\tan(x) \sim x$(当 $x \to 0$). 具体过程可以自行完成, 基本方法是：对第一个积分采用 $1/\tan(x) \sim 1/x$(当 $x \to 0$), 并由极限比较判别法知积分发散; 对第二个积分采用 $1/\sqrt{\tan(x)} \sim 1/\sqrt{x}$(当 $x \to 0^+$), 并由极限比较判别法知该积分收敛.

这是另一个例子：积分

$$\int_0^1 \frac{\sin(x)}{x^{3/2}}\mathrm{d}x.$$

没有因子 $\sin(x)$, 积分根本不会收敛, 因为 3/2 大于 1, 由 p 判别法可知积分发散. 但因子 $\sin(x)$ 改变了这种状况：

$$\frac{\sin(x)}{x^{3/2}} \sim \frac{x}{x^{3/2}} = \frac{1}{x^{1/2}}, \quad 当 x \to 0^+.$$

因为 $\int_0^1 1/x^{1/2}\mathrm{d}x$ 收敛, 由极限比较判别法可知原积分收敛. 这个例子有意思的地方是积分

$$\int_1^\infty \frac{\sin(x)}{x^{3/2}}\mathrm{d}x$$

也收敛, 但原因却完全不同. 这里瑕点在 ∞, 我们要使用绝对积分, 对绝对积分进行直接比较有

$$\int_1^\infty \frac{|\sin(x)|}{x^{3/2}}\mathrm{d}x \leqslant \int_1^\infty \frac{1}{x^{3/2}}\mathrm{d}x < \infty,$$

所以原积分收敛 (这里用了 p 判别法、比较判别法和绝对收敛判别法). 注意在 ∞ 处, 比较好的幂次为 3/2(要是 1/2 就糟了!) 且正弦函数没起任何帮助作用 (也没帮倒忙). 这里我们也顺便得出

$$\int_0^\infty \frac{\sin(x)}{x^{3/2}}\mathrm{d}x$$

收敛, 知道为什么吗?

注: 虽然我们只讨论当 $x \to 0$ 的情况, 但这并不意味着瑕点必须在 0 处, 也可能在 ∞ 处的, 就像下面的例子：

$$\int_1^\infty \sin\left(\frac{1}{x}\right)\mathrm{d}x.$$

这里瑕点在 ∞ 处, 但当 $x \to \infty$ 时 $1/x$ 变得很小. 所以在关系 $\sin(x) \sim x$(当 $x \to 0$) 中, 将 x 换为 $1/x$ 可得当 $1/x \to 0$ 时, $\sin(1/x) \sim 1/x$. 当然, 当 $x \to \infty$ 时 $1/x \to 0$, 所以我们有

$$\sin\left(\frac{1}{x}\right) \sim \frac{1}{x}, \quad 当x \to \infty.$$

现在可由极限比较判别法得积分发散, 因为 $\int_1^\infty 1/x\mathrm{d}x$ 发散.

21.4.3 指数函数在 0 附近的表现

感觉上, **指数函数对 0 没有作用**, 更准确地,

$$当\ x \to 0\ 时, \quad \mathrm{e}^x \sim 1\ 和\ \mathrm{e}^{-x} \sim 1$$

这其实是

$$\lim_{x\to 0} \mathrm{e}^x = 1 \quad 和 \quad \lim_{x\to 0} \mathrm{e}^{-x} = 1$$



的另一种说法. 例如, 反常积分

$$\int_0^1 \frac{\mathrm{e}^x}{x\cos(x)}\mathrm{d}x$$

发散, 因为

$$当\ x \to 0\ 时, \quad \frac{\mathrm{e}^x}{x\cos(x)} \sim \frac{1}{x\cdot 1} = \frac{1}{x}$$



(剩下细节请自行完成.) 注意: 这只对指数 (如 x 或 $-x$) 很小的情况成立. 另一个容易出错的积分是

$$\int_0^1 \frac{\mathrm{e}^{-1/x}}{x^5}\mathrm{d}x.$$

写成 $\mathrm{e}^{-1/x} \sim 1$ 就错了, 因为当 $x \to 0^+$ 时 $1/x \sim \infty$. 我们确实需要采用 21.3.3 节的方法. 特别地, 对任意 n 有

$$\mathrm{e}^{-\text{ 某大量}} \leqslant \frac{C}{(\text{同一大量})^n}.$$

若大的量为 $1/x$(因 x 很小且为正, 所以 $1/x$ 很大), 则变为对任意 n 有

$$\mathrm{e}^{-1/x} \leqslant \frac{C}{(1/x)^n} = Cx^n$$



现在我把证明选择任意大于 4 的 n 命题均成立的任务留给你来完成. 例如, 取 $n = 5$ 可得

$$\int_0^1 \frac{\mathrm{e}^{-1/x}}{x^5}\mathrm{d}x \leqslant \int_0^1 \frac{Cx^5}{x^5}\mathrm{d}x = C\int_0^1 1\mathrm{d}x < \infty,$$



其中最后一个积分由于没有瑕点而显然收敛 (实际上积分值为 1). 顺便说一下, 这是一个相当难的问题.

这是另一个可能的陷阱：在积分

$$\int_0^2 \frac{\mathrm{d}x}{\sqrt{\mathrm{e}^x - 1}}$$

中, 你可能会试图用关系当 $x \to 0$ 时有 $\mathrm{e}^x \sim 1$ 来得出当 $x \to 0$ 时有 $\mathrm{e}^x - 1 \sim 0$, 但后一个关系是错误的, 因为不允许除以 0, 我们需要更聪明点. 在 20.4.1 节, 我们应用了 9.4.2 节中的经典极限

$$\lim_{x \to 0} \frac{\mathrm{e}^x - 1}{x} = 1$$

得到了

$$\boxed{当\ x \to 0, \mathrm{e}^x - 1 \sim x.}$$

据此可得

$$当\ x \to 0^+\ 时, \quad \frac{1}{\sqrt{\mathrm{e}^x - 1}} \sim \frac{1}{\sqrt{x}},$$

现在由极限比较法可知原积分收敛.

21.4.4　对数函数在 0 附近的表现

这里的原理是, 当 $x \to 0^+$ 时对数函数缓慢趋于 $-\infty$. 现在通过取绝对值让对数趋于 ∞, 要知道当 $0 < x < 1$ 时对数值为负, 所以无论 $\alpha > 0$ 有多小, 都存在常数 C 使得

$$\boxed{对于所有\ 0 < x < 1, \quad |\ln(x)| \leqslant \frac{C}{x^\alpha}}$$

这是由 9.4.6 节中的极限 (除了将 a 用 α 代替之外)

$$\lim_{x \to 0^+} x^\alpha \ln(x) = 0$$



推出来的. 这与 21.3.3 节开始采用的论证极为类似.

所以, 为了理解

$$\int_0^1 \frac{|\ln(x)|}{x^{0.9}} \mathrm{d}x,$$

我们采用之前用过多次的方式来讨论这个新的问题. 若没有 $|\ln(x)|$, 积分将收敛. 我们要找一个很小的幂次, 使得它与 0.9 的和仍小于 1. 令 $\alpha = 0.05$ 看一下, 由上面方框中的不等式知有 $|\ln(x)| \leqslant C/x^{0.05}$, 故

$$\frac{|\ln(x)|}{x^{0.9}} \leqslant \frac{C/x^{0.05}}{x^{0.9}} = \frac{C}{x^{0.9}x^{0.05}} = \frac{C}{x^{0.95}}.$$



现在可用比较判别法和 p 判别法来完成该问题, 结果为该积分收敛. 你应该相信若选择任意大于等于 0.1 的数作为 α 的值, 那么就无法得到结论, 又会白费力气了. 顺便说一下, 现在我们自然可知

$$\int_0^1 \frac{\ln(x)}{x^{0.9}}\mathrm{d}x$$

收敛, 因为它是原积分求负得来的.

考虑另外一个例子

$$\int_0^{1/2} \frac{1}{x^2|\ln(x)|}\mathrm{d}x.$$

若没有因子 $|\ln(x)|$, 由 p 判别法可知积分发散. $|\ln(x)|$ 有使积分收敛的趋势, 但作用不大, 因为它只是对数, 而对数增长缓慢. 所以我们仍预期积分发散. 为了证明该猜测, 注意 $|\ln(x)| \leqslant C/x^\alpha$, 取倒数可得 $1/|\ln(x)| \geqslant x^\alpha/C$. 为了避免徒劳无功, 我们再一次选择足够小的 α, 有

$$\frac{1}{x^2|\ln(x)|} \geqslant \frac{x^\alpha}{Cx^2},$$

所以只要 $\alpha \leqslant 1$ 就可以. (为什么？) 实际上, 当 $\alpha = 1$ 时右边变为 $1/(Cx)$, 到这里就可知积分发散. 注意积分

$$\int_0^{1/2} \frac{1}{x^2\ln(x)}\mathrm{d}x$$

也发散 (趋于 ∞), 因为它是原积分求负的结果.

最后一个例子：积分

$$\int_0^{1/2} \frac{1}{x^{0.9}|\ln(x)|}\mathrm{d}x.$$

现在积分在没有因式 $|\ln(x)|$ 时收敛, 但将这个很大的量放到分母上只会使积分收敛得更快, 所以只需找到 $|\ln(x)|$ 在 $(0,\ 1/2]$ 的最小值. 想一想并确定当 $x = 1/2$ 时有最小值, 所以当 $0 < x \leqslant 1/2$, 我们有 $|\ln(x)| \geqslant |\ln(1/2)| = \ln(2)$. 最后, 两边取倒数并除以 $x^{0.9}$ 可得对所有 $0 < x \leqslant 1/2$, 有

$$\frac{1}{x^{0.9}|\ln(x)|} \leqslant \frac{1}{x^{0.9}\ln(2)}$$

成立. 现在只需运用比较判别法和 p 判别法可得原积分收敛.

21.4.5 更一般的函数在 0 附近的表现

在 24.2.2 节, 我们将学习麦克劳林级数. 如果之前没见过, 不要着急! 留下标记, 等学完麦克劳林级数的所有内容后再来读本节. 不管怎样, 基本观点是：若一个函数有在 0 附近收敛于该函数的麦克劳林级数, 则函数在 $x \to 0$ 时渐近等价于级数的最低次项, 即

$$\boxed{f(x) = a_n x^n + a_{n+1}x^{n+1} + \cdots, \text{当 } x \to 0, \text{ 则} f(x) \sim a_n x^n.}$$

考虑下面的例子：

$$\int_0^1 \frac{\mathrm{d}x}{1-\cos(x)} \quad 和 \quad \int_0^1 \frac{\mathrm{d}x}{(1-\cos(x))^{1/3}}.$$

我们知道当 $x \to 0$ 时 $\cos(x) \sim 1$, 但这并没有告诉我们 $1-\cos(x)$ 怎样. 讨论这个量的一个方法是运用 $\cos(x)$ 的麦克劳林级数:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots;$$

它可以另写为

$$1 - \cos(x) = \frac{x^2}{2} - \frac{x^4}{24} + \cdots.$$

所以, 由上面的原理知右边最低次项起决定作用, 我们有

$$当\ x \to 0\ 时, \quad 1-\cos(x) \sim \frac{x^2}{2}.$$

这与我们在 7.1.2 节讨论的例子

$$\lim_{x\to 0} \frac{1-\cos(x)}{x^2} = \frac{1}{2}$$



一致. 我把利用渐近等价关系证明上面第一个积分发散, 第二个收敛的任务留作练习.

21.5 如何应对不在 0 或 ∞ 处的瑕点



若瑕点出现在有限值而非 0 处, 做换元. 具体情况如下.

- 若积分 $\int_a^b f(x)\mathrm{d}x$ 的唯一瑕点出现在 $x = a$ 处, 做换元 $t = x - a$, 注意 $\mathrm{d}t = \mathrm{d}x$. 新的积分则只有 0 一个瑕点.



- 若积分 $\int_a^b f(x)\mathrm{d}x$ 的唯一瑕点出现在 $x = b$ 处, 做换元 $t = b - x$, 注意 $\mathrm{d}t = -\mathrm{d}x$, 用多出的负号来做积分上下限交换. 新的积分则只有 0 一个瑕点.

例如, 我们在 20.1.2 节讨论了

$$\int_0^3 \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x.$$

我们将该积分拆分成了 5 个积分, 每个积分只有一个瑕点, 并证明了它们均发散. 其中一个积分 (我们称之为 I_5) 为

$$\int_2^3 \frac{1}{x(x-1)(x+1)(x-2)}\mathrm{d}x.$$

这里瑕点在 $x=2$ 处, 故做换元 $t = x-2$. 由此 $x = t+2$, 积分变为

$$\int_0^1 \frac{1}{(t+2)(t+1)(t+3)t}\mathrm{d}t.$$

积分的上下限现在为 1 和 0, 瑕点变为 0. 现在我们可运用多项式的最低次项在 0 附近起决定作用的事实得

$$\text{当 } t \to 0 \text{ 时}, \quad t+2 \sim 2, \quad t+1 \sim 1, \quad t+3 \sim 3.$$

综合以上事实可知

$$\text{当 } t \to 0 \text{ 时}, \quad \frac{1}{(t+2)(t+1)(t+3)t} \sim \frac{1}{2 \times 1 \times 3 \times t} = \frac{1}{6t}.$$

由极限比较判别法和 p 判别法知上述积分发散.

另一个由原积分拆出的积分 (我们称之为 I_4) 为

$$\int_{3/2}^{2} \frac{1}{x(x-1)(x+1)(x-2)} \mathrm{d}x.$$

现在瑕点在 $x = 2$ 处, 是积分的右极限. 故做换元 $t = 2 - x$. 当 $x = 3/2$ 时有 $t = 1/2$, 且当 $x = 2$ 时 $t = 0$. 由 $\mathrm{d}t = -\mathrm{d}x$ 和 $x = 2 - t$, 我们有

$$\begin{aligned}\int_{3/2}^{2} \frac{1}{x(x-1)(x+1)(x-2)} \mathrm{d}x &= -\int_{1/2}^{0} \frac{1}{(2-t)(1-t)(3-t)(-t)} \mathrm{d}t \\ &= \int_{0}^{1/2} \frac{1}{(2-t)(1-t)(3-t)(-t)} \mathrm{d}t.\end{aligned}$$

在最后一个积分中, 我们用等式 $\mathrm{d}x = -\mathrm{d}t$ 中的负号来交换积分的上下限 (如 16.3 节所述). 总之, 很容易得到

$$\text{当 } t \to 0 \text{ 时}, \quad \frac{1}{(2-t)(1-t)(3-t)(-t)} \sim -\frac{1}{6t}.$$

所以上述积分发散 (还是根据极限比较判别法和 p 判别法, 细节自行完成, 处理被积函数的负号时要小心). 现在, 你就可以试着证明其他三个积分 (20.1.2 节的 I_1、I_2 和 I_3) 发散了.



第 22 章　数列和级数：基本概念

无穷级数和反常积分非常相似, 这是个好消息. 所以, 很多 (但不是全部) 反常积分的方法都可以用于讨论无穷级数, 我们就不用重新寻找方法了. 要定义无穷级数, 就要先讨论数列. 跟反常积分的讨论一样, 我用两章来讨论数列和级数：本章主要包括一些原理, 而下一章注重实际, 包含了求解问题的若干方法. 如果你是第一次阅读, 那就先看一下本章的详细内容吧. 如果是为了回顾, 快速浏览一下要点就足够了, 然后可以直接看下一章的具体例题. 下面是本章的内容：

- 数列的收敛和发散;
- 两个重要数列;
- 数列极限和函数极限之间的联系;
- 级数的收敛与发散, 以及几何级数的敛散性讨论;
- 级数的第 n 项判别法;
- 级数和反常积分的联系;
- 比式判别法、根式判别法、积分判别法以及交错级数判别法的介绍.

本章主要进行理论探讨, 大部分例题在下一章.

22.1　数列的收敛和发散

数列是一列有序的数, 可能有有限项, 也可能有无穷项, 其中有无穷项的数列叫作*无穷数列*. 例如,

$$0, 1, -1, 2, -2, 3, -3, \cdots$$

是一个包含所有整数的无穷数列. 下角标经常用于数列中, 其中 a_1 表示数列中的第一项, a_2 表示第二项, a_3 表示第三项, 以此类推. (有时 a_0 是第一项, a_1 是第二项, 以此类推. 我们也可以不用 a, 如用 b_n 或其他的字母.) 所以上例中, $a_1 = 0$、$a_2 = 1$、$a_3 = -1$、$a_4 = 2$, 以此类推. 数列经常由一个公式来给出, 如

$$a_n = \frac{\sin(n)}{n^2},$$

其中 $n = 1, 2, \cdots$, 定义了数列

$$\frac{\sin(1)}{1^2}, \quad \frac{\sin(2)}{2^2}, \quad \frac{\sin(3)}{3^2}, \quad \frac{\sin(4)}{4^2}, \cdots$$

对于无穷数列, 我们主要讨论当 n 趋于无穷时数列的极限值, 即当我们观察数列中越来越靠后的数时, 会发生什么？数学上表示为, 极限

$$\lim_{n\to\infty} a_n$$

存在与否; 若存在, 值是多少. 虽然我们还未给出上述极限的定义, 不过它与函数 f 的极限 $\lim_{x\to\infty} f(x)$ 差不多. (定义参见附录 A 的 A.3.3 节.) 基本思想是:

$$\lim_{n\to\infty} a_n = L$$

意味着 a_n 在开始时可能有稍许徘徊, 最后会越来越趋近于 L 并一直保持这种趋势. 若存在这样的 L, 则数列 $\{a_n\}$ **收敛**, 否则**发散**. 与函数一样, 数列也可以发散到 ∞ 或 $-\infty$, 也可以不断振荡 (可能会很疯狂) 而不趋于一个特定的值. 例如, 上述数列 $0, 1, -1, 2, -2, \cdots$ 发散, 但不是发散到 ∞ 或 $-\infty$, 而是在绝对值不断增大的正数和负数间振荡.

和函数一样, 有时也可以说当 $n \to \infty$ 时 $a_n \to L$, 这与 $\lim_{n\to\infty} a_n = L$ 意思一样.

22.1.1 数列和函数的联系

考虑数列

$$a_n = \frac{\sin(n)}{n^2},$$

我们之前见过, 它与函数

$$f(x) = \frac{\sin(x)}{x^2}$$

紧密相关. 事实上, 对每个正整数 n, a_n 都等于 $f(n)$. 所以, 如果我们能证明 $\lim_{x\to\infty} f(x)$ 存在, 就可以说数列 $\{a_n\}$ 有相同的极限. 数列继承了函数的极限性质. 在水平渐近线上, 二者也有联系: 记住, 若 $\lim_{x\to\infty} f(x) = L$, 则 $y = f(x)$ 的图像有水平渐近线 $y = L$.

除了上述讨论外, 我们还可以很容易地将函数极限的其他性质推广到数列极限. 例如两个收敛数列 $\{a_n\}$ 和 $\{b_n\}$, 当 $n \to \infty$ 时, $a_n \to L$, $b_n \to M$, 则其和 $a_n + b_n$ 构成一个收敛于 $L + M$ 的新数列. 对于差、积、商 (假定 $M \neq 0$, 因为分母不能为 0) 和常数的积也同样适用. 虽然这个结论意义没有那么深远, 不过的确很有用.

另一个重要的事实是三明治定理, 即夹逼定理, 对数列也适用. (三明治定理内容参见 3.6 节.) 特别地, 假设有数列 $\{a_n\}$, 若怀疑其收敛于某数 L, 则要找到一个比 $\{a_n\}$ 大的数列 $\{b_n\}$ 和一个比其小的数列 $\{c_n\}$, 且两个数列均收敛于 L, 则我们就可知该数列的确收敛于 L 了. 用数学语言描述就是, 若 $c_n \leqslant a_n \leqslant b_n$, 且当 $n \to \infty$ 时, $b_n \to L$, $c_n \to L$, 则当 $n \to \infty$ 时, $a_n \to L$. 对前面的数列



$$a_n = \frac{\sin(n)}{n^2}$$

可以通过将经典不等式 $-1 \leqslant \sin(n) \leqslant 1$ 除以 n^2, 并利用三明治定理得对所有 n 有

$$\frac{-1}{n^2} \leqslant \frac{\sin(n)}{n^2} \leqslant \frac{1}{n^2}.$$

数列 $b_n = 1/n^2$ 和 $c_n = -1/n^2$ 在 $n \to \infty$ 时均收敛于 0, 所以夹于它们之间的数列 a_n 也收敛于 0. 即

$$\lim_{n\to\infty} \frac{\sin(n)}{n^2} = 0.$$



另一个可由函数性质推广过来的是**连续函数保持极限**. 这是什么意思呢? 假设当 $n \to \infty$ 时 $a_n \to L$, 则如果函数 f 在 $x = L$ 连续, 我们就可以说当 $n \to \infty$ 时 $f(a_n) \to f(L)$. 当对任何式子取函数 f 时, 极限关系仍保持. 例如求



$$\lim_{n\to\infty} \cos\left(\frac{\sin(n)}{n^2}\right)$$

是多少? 我们已经有

$$\text{当 } n \to \infty \text{ 时}, \quad \frac{\sin(n)}{n^2} \to 0,$$

由于余弦函数在 0 点连续, 因而两边同时取余弦, 可得

$$\text{当 } n \to \infty \text{ 时}, \quad \cos\left(\frac{\sin(n)}{n^2}\right) \to \cos(0) = 1.$$



还有一个可以从函数理论中借用的重要工具是洛必达法则(见 14.1 节). 应用该法则的一个问题是, 不能对关于 n 的量 a_n 求导, 因为 n 只是一个整数. 事实上, 当对函数 f 求关于变量 x 的导数时, 只是为了看一下当对 x 做极小变动时函数 $f(x)$ 有什么变化. 你不能对整数做极小变动, 因为极小变动后它就不再是整数了. 所以若想应用洛必达法则, 首先需将数列嵌入到一个合适的函数中. 例如, 若 $a_n = \ln(n)/\sqrt{n}$, 则可令

$$f(x) = \frac{\ln(x)}{\sqrt{x}},$$

并利用洛必达法则求出 $\lim\limits_{x\to\infty} f(x)$ 的值再求得 $\lim\limits_{n\to\infty} a_n$. 注意, 这是 ∞/∞ 情形, 所以可以利用该法则. 对分子和分母分别求导, 可得

$$\lim_{x\to\infty} \frac{\ln(x)}{\sqrt{x}} \overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{1/x}{1/2\sqrt{x}} = \lim_{x\to\infty} \frac{2}{\sqrt{x}} = 0.$$

因为函数的极限是 0, 则数列 a_n 当 $n \to \infty$ 时也收敛于 0.(我们也可以采用对数在 ∞ 处增长缓慢的结论来求上述极限, 只需要应用 21.3.4 节开头部分的公式并令 $\alpha = 1/2$ 即可.)

22.1.2　两个重要数列

取常数 r, 并考虑从 $n = 0$ 开始取值的数列 $a_n = r^n$, 这是一个等比数列, 每一项都是前一项与这个常数的乘积. 我们来看一些等比数列:

- 若 $r = 0$, 则数列为 $0, 0, 0, \cdots$, 显然收敛于 0;

- 若 $r = 1$, 则数列为 $1, 1, 1, \cdots$, 显然收敛于 1;
- 若 $r = 2$, 则数列为 $1, 2, 4, 8, \cdots$, 明显发散于 ∞;
- 若 $r = -1$, 则数列为 $1, -1, 1, -1, 1, \cdots$, 发散, 但不是发散于 ∞ 或 $-\infty$, 因为它一直在 -1 和 1 之间来回振荡, 换句话说, 不存在极限;
- 若 $r = -2$, 则数列为 $1, -2, 4, -8, \cdots$, 与上面数列发散方式相同 (不存在极限), 事实上这次的振荡范围更宽;
- 若 $r = 1/2$, 则数列为 $1, 1/2, 1/4, 1/8, \cdots$, 收敛于 0;
- 若 $r = -1/2$, 则数列为 $1, -1/2, 1/4, -1/8, \cdots$, 尽管振荡, 也收敛于 0, 因为振荡最后变得越来越小.

上面这些都是下述一般规则的特例.

$$\lim_{n\to\infty} r^n \begin{cases} = 0 & \text{如果 } -1 < r < 1, \\ = 1 & \text{如果 } r = 1, \\ = \infty & \text{如果 } r > 1, \\ \text{不存在} & \text{如果 } r \leqslant -1. \end{cases}$$

我们对上述极限进行证明. 首先, 当 $r \geqslant 0$, 极限与 9.4.4 节 (见中间的方框) 中的含有 r^x 的极限相似. 容易出错的情况是当 $r < 0$ 时, 这是因为数列振荡. 为了解决这个问题, 注意对所有 n 都有

$$-|r|^n \leqslant r^n \leqslant |r|^n,$$

这里比较好的情况是数列 $\{-|r|^n\}$ 和 $\{|r|^n\}$ 都不振荡. 实际上, 若 $-1 < r < 0$, 则 $|r| < 1$, 因此我们知道这两个数列都收敛于 0, 现在可用三明治定理推出 $r^n \to 0$. 最后, 若 $r \leqslant -1$, 则 r^n 不可能收敛, 因为它的值在大于等于 1 和小于等于 -1 的数中来回跳跃, 则极限因这些振荡而不存在. (该情形与 3.4 节的极限 $\lim\limits_{x\to\infty} \sin(x)$ 类似, 也可参见附录 A 的 A.3.4 节.)

等比数列无须从 1 开始, 若令 $a_n = ar^n$, 其中 a 为常数, 则首项 a_0 等于 a. 你可以将上述方框中的 $\lim\limits_{n\to\infty} r^n$ 的值乘以 a 来求 $\lim\limits_{n\to\infty} ar^n$ 的值. 最重要的是, 若 $-1 < r < 1$, 则 $\lim\limits_{n\to\infty} ar^n$ 为 0, 与 a 无关.

把大量时间用在等比数列的讨论之后, 我们来快速看另一个数列. 特别地, 若 k 为任意常数, 则

$$\lim_{n\to\infty} \left(1 + \frac{k}{n}\right)^n = \mathrm{e}^k.$$

这就是根据 9.2.3 节开头讲的那个极限而来的. 在数列的相关内容中可知, 这个极限很有用.

22.2 级数的收敛与发散

级数就是和, 就是将数列 a_n 的所有项都加起来, 把各项之间的逗号用加号代替. 对于无穷数列, 好像有点理不出头绪了, 将无穷多个数相加意味着什么呢? 例如, 若数列 a_n 是等比数列 $1, 1/2, 1/4, 1/8, \cdots$, 则相应的级数就是 $1+1/2+1/4+1/8+\cdots$. 我们需要做一些不同寻常的事情, 来处理意味着级数不断加下去的省略号.

一般地, 我们想知道

$$a_1 + a_2 + a_3 + \cdots$$

意味着什么. 为了处理这个无穷项之和, 我们把前若干项之后的项去掉. 若取前面 N 项, 则去掉后面那些项的级数为

$$a_1 + a_2 + a_3 + \cdots + a_{N-1} + a_N.$$

现在是有限项之和, 变得有意义了. 下面就是我们想要的:

$$a_1 + a_2 + a_3 + \cdots = \lim_{N\to\infty} (a_1 + a_2 + a_3 + \cdots + a_{N-1} + a_N).$$

右边看起来有点奇怪, 因为随着 N 的增大, 项数也在增多. 所以, 我们定义一个新的数列 $\{A_N\}$:

$$A_N = a_1 + a_2 + a_3 + \cdots + a_{N-1} + a_N.$$

这个新的数列被称为*部分和数列*. 前面那个奇怪的等式现在为

$$a_1 + a_2 + a_3 + \cdots = \lim_{N\to\infty} A_N.$$

现在右边就不那么奇怪了, 是个数列的极限. 如果极限存在且等于 L, 则我们说左边的级数收敛于 L. 若极限不存在, 则级数发散.

理解上面这些, 有个极好的类比. 假想你站在一条又直又长的高速公路休息站旁, 休息站两边公路向两侧延伸, 一边是来的方向, 另一边是要去的方向. 休息站的位置为 0. (我们在 5.2.2 节见过这个高速公路的例子.) 不幸的是, 你失去了自由意识, 有人每分钟都用扩音器告诉你走一定的英尺数, 他下命令你才能动. 如果他说出一个负数, 你就往回走, 每一次移动称为一步. (希望他不会让你一步走 100 英尺!)

持扩音器的家伙喊的第一个数为 a_1, 你从位置 0 移动到位置 a_1 (长度单位是英尺, 下同). 第二个数为 a_2, 又向前走 a_2 英尺, 现在到哪儿了? 在位置 $a_1 + a_2$ 处, 因这次是从 a_1 开始走的. 在第三个数 a_3 之后, 你将在位置 $a_1 + a_2 + a_3$. 趋势很明显: 按 a_1, a_2, a_3 的步长, 第 N 步为 a_N, 你会在位置

$$a_1 + a_2 + a_3 + \cdots + a_{N-1} + a_N.$$

这正好是上面定义的部分和 A_N 的值. 换句话说, A_N 是第 N 步后你的位置. 所以当有

$$a_1 + a_2 + a_3 + \cdots = \lim_{N\to\infty} A_N,$$

意思就是如果最终要走向高速公路的某个特定目标, 你可以将所有的步都加起来. 你必须非常非常接近那个点, 决不能很远. 在那个点附近, 要用很小的步子踮起脚走, 否则就不可能把这些步加起来, 级数将会发散.

现在是时候引入求和号了 (见 15.1 节). A_N 表达式变为

$$A_N = a_1 + a_2 + a_3 + \cdots + a_{N-1} + a_N = \sum_{n=1}^{N} a_n.$$

无穷级数可写为

$$a_1 + a_2 + a_3 + \cdots = \sum_{n=1}^{\infty} a_n.$$

因此, 下面是用求和号定义的无穷级数的值:

$$\boxed{\sum_{n=1}^{\infty} a_n = \lim_{N\to\infty} \sum_{n=1}^{N} a_n.}$$

如果右边的极限不存在, 则左边的级数发散. 右边就是数列的极限, 所以上述等式并不像符号所表示的那样简易明了.

我们再回顾一下. 我们从无穷**数列**

$$\{a_n\} = a_1, a_2, a_3, \cdots$$

开始, 并用它构造无穷**级数**

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots.$$

为了理解级数的极限, 这里构造了一个新的部分和**数列**:

$$A_N = \sum_{n=1}^{N} a_n = a_1 + a_2 + a_3 + \cdots + a_{N-1} + a_N.$$

由定义, 如果极限存在, 级数的极限与部分和数列的极限一样; 否则级数发散. 鉴于这里有两个数列与一个级数一起讨论, 一定要确保能将它们区分清楚!

级数不必从 $n=1$ 开始, 也可以从其他的数开始, 甚至 $n=0$. 你需要做的仅仅是把部分和的起始项更改一下. 重要的一点是: 级数收敛还是发散与起始项无关! 例如, 我们将在 22.4.3 节看到级数



$$\sum_{n=1}^{\infty} \frac{1}{n}$$

发散. 由此结果, 马上可知下面的级数也发散:

$$\sum_{n=5}^{\infty} \frac{1}{n}, \quad \sum_{n=89}^{\infty} \frac{1}{n}, \quad \sum_{n=1\,000\,000}^{\infty} \frac{1}{n}.$$

为了讨论为什么第一个级数发散, 只要将原来那个级数的前四项取出来, 如下:

$$\sum_{n=1}^{\infty}\frac{1}{n}=\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\sum_{n=5}^{\infty}\frac{1}{n}=\frac{25}{12}+\sum_{n=5}^{\infty}\frac{1}{n}.$$

从 $n=1$ 开始的级数和从 $n=5$ 开始的级数只相差有限常数 25/12. 由于从 $n=1$ 开始的级数发散到 ∞, 故减去 25/12 对其不会有任何影响, 因而从 $n=5$ 开始的级数一定也发散. 当然, 5 没有什么特别的, 对任意的起始点都会有相同的结果. 类似的, 我们将在 22.4.3 节讨论

$$\sum_{n=1}^{\infty}\frac{1}{n^2}$$



实际上也是收敛的. 这意味着如果能将原和式进行拆分并证明的话, 下面的这些级数也收敛：

$$\sum_{n=4}^{\infty}\frac{1}{n^2},\quad \sum_{n=101}^{\infty}\frac{1}{n^2},\quad \sum_{n=5\,000\,000}^{\infty}\frac{1}{n^2}.$$

在我们讨论几何级数之前再注意一件事：考虑

$$\sum_{n=0}^{\infty}\frac{1}{n^2}.$$

我们将起始点换为 $n=0$, 令人讨厌的事是：首项变为 $1/0^2$, 但它并不存在. 因此上述级数不是发散, 而是没意义, 因为首项没有定义. 我们总是以一个足够大的 n 作为起点以避免这样的情形, 这样级数的所有项就都有定义了.

几何级数(理论)

我们来看一个无穷级数的重要例子. 假定以我们在 22.1.2 节见过的等比数列 $1, r, r^2, r^3, \cdots$ 开始, 可以把这个数列作为无穷级数

$$1+r+r^2+r^3+\cdots=\sum_{n=0}^{\infty}r^n$$

的项, 则这个级数为几何级数. 问题是, 该级数收敛吗？若收敛, 收敛于何值？

为了求解, 我们最好看一下部分和. 选择数 N, 则部分和 A_N 为

$$A_N=1+r+r^2+r^3+\cdots+r^{N-1}+r^N.$$

用求和号表示为

$$A_N=\sum_{n=0}^{N}r^n.$$

希望你在前面的学习中已经知道上述表达式可化简为

$$A_N=1+r+r^2+r^3+\cdots+r^{N-1}+r^N=\frac{1-r^{N+1}}{1-r},$$

只要 $r \neq 1$. (不管怎样, 后面会给出其证明.) 现在我们要求当 $N \to \infty$ 时 A_N 的极限, 首先假设 $-1 < r < 1$, 则由前面 22.1.2 节方框中第一种情形知 $\lim\limits_{N\to\infty} r^N = 0$, 将 N 换为 $N+1$ 也得到 $\lim\limits_{N\to\infty} r^{N+1} = 0$. 所以

$$\lim_{N\to\infty} A_N = \lim_{N\to\infty} \frac{1-r^{N+1}}{1-r} = \frac{1}{1-r}.$$

该几何级数收敛于 $1/(1-r)$. 下面是写在一起的带求和号的整个论证过程：

$$\sum_{n=0}^{\infty} r^n = \lim_{N\to\infty} \sum_{n=0}^{N} r^n = \lim_{N\to\infty} \frac{1-r^{N+1}}{1-r} = \frac{1}{1-r}.$$

若 r 不介于 1 和 -1 之间呢? 结论是几何级数肯定发散, 下一节将给出证明. 总结如下：

> 如果 $-1 < r < 1$, $\quad \displaystyle\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$;
>
> 如果 $r \geqslant 1$ 或 $r \leqslant -1$, 级数发散.

上述几何级数的首项总是 1, 因为 $r^0 = 1$. 如果用其他的某数 a 代替, 则各项为 a, ar, ar^2 等. 所以每项都可以乘 a, 得到上述原理更一般的形式：

> 如果 $-1 < r < 1$, $\quad \displaystyle\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$;
>
> 如果 $r \geqslant 1$ 或 $r \leqslant -1$, 级数发散.

我们将在 23.1 节讨论几何级数的更多例子. 同时, 证明

$$A_N = \sum_{n=0}^{N} r^n = \frac{1-r^{N+1}}{1-r}.$$

证明如下：首先, 和式左乘 $(1-r)$ 可得

$$A_N(1-r) = (1-r)\sum_{n=0}^{N} r^n.$$

将因式 $(1-r)$ 移入求和号内并化简得

$$A_N(1-r) = \sum_{n=0}^{N} r^n(1-r) = \sum_{n=0}^{N} (r^n - r^{n+1}).$$

右边的和是一个伸缩级数 (见 15.1.2 节), 所以和为 $r^0 - r^{N+1}$ 或 $1 - r^{N+1}$. 因此 $A_N(1-r) = 1 - r^{N+1}$. 现在为证得结果, 只需除以 $(1-r)$, 其中 $(1-r)$ 不为 0, 因为我们已经假设 $r \neq 1$.

22.3 第 n 项判别法 (理论)

对于收敛级数, 部分和的极限必须存在. 要知道, N 步后的部分和表示你按照持扩音器家伙的指令走了 N 步后的位置. (若不明白我在说什么, 参见 22.2 节) 总之, 若你的位置随着你的步数不断增加而逐渐收敛于某个极限位置, 则每一步都需要变得很小很小, 否则, 你将失误且不能待在与特定位置一致的地方. 前后挪动并不好, 要接近特定位置, 你需要非常靠近, 驻留在很接近的位置.

所以, 由数列 $\{a_n\}$ 给出的每一步到最后要变得很小很小, 才能使得级数收敛. 数学上表示为, 当 $n \to \infty$ 时 $a_n \to 0$, 故我们有

第 n 项判别法: 若 $\lim\limits_{n\to\infty} a_n \neq 0$, 或极限不存在, 则级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

若 $\lim\limits_{n\to\infty} a_n = 0$, 则级数可能收敛也可能发散, 需要采用其他的方法解决该问题. 注意: **第 n 项判别法不能用于级数收敛性的判别!**

这个判别法是一种求真判定: 若 a_n 不趋于 0, 该级数发散. 否则仍需要采用其他方法继续讨论该问题. 例如, 我们马上要讨论

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \text{ 收敛, 但 } \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \text{ 发散.}$$

这两个级数的通项都趋于 0:

$$\lim_{n\to\infty} \frac{1}{n^2} = 0, \quad \lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0.$$

第 n 项判别法在两个级数中都不适用! 只有当极限不为 0 的时候才能使用该判别法. 下面是一些该判别法适用的例子:

$$\sum_{n=0}^{\infty} 2^n, \quad \sum_{n=0}^{\infty} (-3)^n \text{ 和 } \sum_{n=0}^{\infty} 1.$$

我们有

$$\lim_{n\to\infty} 2^n = \infty, \quad \lim_{n\to\infty} (-3)^n \text{ 不存在, } \lim_{n\to\infty} 1 = 1.$$

根据第 n 项判别法, 上面三个级数都发散, 因为每个级数通项的极限都不是 0. 事实上, 这些级数都是几何级数, 公比分别为 $2, -3$ 和 1. 一般地, 对于公比 $r \geqslant 1$ 或 $r \leqslant -1$ 的几何级数 $\sum\limits_{n=0}^{\infty} r^n$, 通项当 $n \to \infty$ 时不趋于 0. (见 22.1.2 节方框中公式.) 所以第 n 项判别法告诉我们, 任何公比不在 $(-1, 1)$ 里的几何级数均发散.

在收敛级数中, 虽然通项 a_n 要收敛于 0, 但这并不意味着级数的极限为 0. 例如, 公比为 $r = 1/2$ 的等比数列 $1, 1/2, 1/4, 1/8, \cdots$ 收敛于 0, 我们可以由前一节的

公式得出相应的级数的值：

$$\sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n=\frac{1}{1-r}=\frac{1}{1-\frac{1}{2}}=2.$$

所以数列收敛于 0, 而级数却收敛于 2. 反过来可以说, 若数列收敛于 2, 则由第 n 项判别法知相应的级数发散.

我们将在 23.2 节看到更多关于第 n 项判别法的例子. 现在来看一下其他的判别法.

22.4 无穷级数和反常积分的性质

无穷级数和反常积分之间是有一些联系的, 特别是当反常积分在 ∞ 有瑕点的时候. 其中的一个联系就是积分判别法, 这将在 22.5.3 节讨论. 本节主要告诉你反常积分的四个判别法对无穷级数仍适用. 下面就一一讨论.

22.4.1 比较判别法 (理论)

假设给定的级数 $\sum\limits_{n=1}^{\infty} a_n$ 的每一项都为正, 若认为该级数发散, 则要能找到一个比 $\sum\limits_{n=1}^{\infty} a_n$ 小的发散级数 $\sum\limits_{n=1}^{\infty} b_n$, 即可被证实. 即, 若对所有 n, 有 $0 \leqslant b_n \leqslant a_n$, 且 $\sum\limits_{n=1}^{\infty} b_n$ 发散, 则 $\sum\limits_{n=1}^{\infty} a_n$ 也发散. 若认为原级数收敛, 则要能找一个比它大的收敛级数 $\sum\limits_{n=1}^{\infty} b_n$, 即可被证实. 即, 若对所有 n, 有 $b_n \geqslant a_n \geqslant 0$, 且 $\sum\limits_{n=1}^{\infty} b_n$ 收敛, 则 $\sum\limits_{n=1}^{\infty} a_n$ 也收敛.

基本上, 这与反常积分的比较判别法一致. 级数比较判别法成立理由与积分的比较判别法一样, 若有兴趣可自行证明.



级数的首项不必从 $n=1$ 开始, 可以从任何数开始. 例如, 考虑

$$\sum_{n=3}^{\infty}\left(\frac{1}{2}\right)^n|\sin(n)|.$$

利用比较判别法很容易判定. 要知道, 对任何 n, 都有 $|\sin(n)| \leqslant 1$, 我们可得

$$\sum_{n=3}^{\infty}\left(\frac{1}{2}\right)^n|\sin(n)| \leqslant \sum_{n=3}^{\infty}\left(\frac{1}{2}\right)^n < \infty.$$

后面一个级数收敛, 因为它是公比为 1/2(介于 -1 到 1 之间) 的几何级数. 根据比较判别法, 原级数也收敛. 下一章我们将介绍比较判别法的更多例子.

22.4.2　极限比较判别法 (理论)

在 20.4.1 节, 我们有如下的定义:

$$\text{当 } x \to \infty \text{ 时,}\quad f(x) \sim g(x) \text{ 与 } \lim_{x\to\infty} \frac{f(x)}{g(x)} = 1 \text{ 含义一样.}$$

这里数列也有类似的定义:

$$\boxed{\text{当 } n \to \infty \text{ 时, } a_n \sim b_n \text{ 与 } \lim_{n\to\infty} \frac{a_n}{b_n} = 1 \text{ 含义一样.}}$$

极限比较判别法为, 若当 $n \to \infty$ 时 $a_n \sim b_n$, 且 a_n 和 b_n 均有限, 则 $\sum\limits_{n=1}^{\infty} a_n$ 与 $\sum\limits_{n=1}^{\infty} b_n$ 同时收敛或同时发散. 当然, 没有必要必须从 $n = 1$ 开始, 也可以从 $n = 0$, $n = 19$ 或任何其他的 n 的有限值开始. 该比较判别法的证明与积分的极限比较法证明类似, 这里不再赘述. 读者可自行证明, 若当 $n \to \infty$ 时 $a_n \sim b_n$, 我们说两个数列渐近等价.

我们在第 21 章讨论的函数的所有性质对数列均成立. 例如, 考虑

$$\sum_{n=0}^{\infty} \sin\left(\frac{1}{2^n}\right).$$

当 n 很大时, $1/2^n$ 变得很小 (即趋于 0). 我们知道当 $x \to 0$ 时 $\sin(x) \sim x$ (见 21.4.2 节), 将 x 用 $1/2^n$ 代换, 我们有

$$\text{当 } \frac{1}{2^n} \to 0 \text{ 时,}\quad \sin\left(\frac{1}{2^n}\right) \sim \frac{1}{2^n}.$$

将 $1/2^n$ 另写为 $(1/2)^n$, 且注意 $1/2^n \to 0$ 等价于 $n \to \infty$. 故上述关系可写成

$$\text{当 } n \to \infty \text{ 时,}\quad \sin\left(\frac{1}{2^n}\right) \sim \left(\frac{1}{2}\right)^n.$$

由极限比较法, 两级数

$$\sum_{n=0}^{\infty} \sin\left(\frac{1}{2^n}\right) \text{ 和 } \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n$$

同时收敛或同时发散. 现在我们知道右边的级数收敛, 因为它是公比为 1/2(绝对值小于 1) 的几何级数, 所以左边的级数也收敛. 但是, 右边的级数收敛于 2(如 22.3 节所见) 并不意味着左边级数也收敛于 2. 我们不知道它收敛于何值, 只可知其收敛.

22.4.3　p 判别法 (理论)

级数也有 p 判别法, 基本上与反常积分在瑕点 ∞ 的 p 判别法一样. 特别地,

$$\sum_{n=a}^{\infty}\frac{1}{n^p}\begin{cases}\text{收敛, 若 } p>1;\\ \text{发散, 若 } p\leqslant 1.\end{cases}$$

它的最简单的证明要用到积分判别法, 所以将到 22.5.3 节再讨论. p 判别法的一些简单例子为



$$\sum_{n=1}^{\infty}\frac{1}{n^2}\text{收敛, 而 }\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\text{发散.}$$

第一个级数中幂次 2 大于 1, 故收敛. 另一方面, 由 $\sqrt{n}=n^{1/2}$ 知第二个级数幂次为 1/2, 因为 1/2 小于 1, 级数发散.

在讨论绝对收敛判别法之前, 先看一下所谓的**调和级数**

$$\sum_{n=1}^{\infty}\frac{1}{n}.$$

由 p 判别法知该级数发散, 不过我们也可以直接证明它发散. 方法是: 先将该级数各项写出来, 再把它们用特殊的方式组合. 特别地, 上述级数可以写为

$$1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)\\+\left(\frac{1}{9}+\frac{1}{10}+\frac{1}{11}+\frac{1}{12}+\frac{1}{13}+\frac{1}{14}+\frac{1}{15}+\frac{1}{16}\right)+\cdots.$$

除了开始的 1 和 1/2 外, 后面每一组中的项数都是前一组的两倍. 主要过程为: 每一组的最后一项是该组中最小的项, 故上面的和式大于

$$1+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)\\+\left(\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}\right)+\cdots.$$

在这个新级数中, 一项为 1, 一项为 1/2, 两项为 1/4, 四项为 1/8, 八项为 1/16, 以此类推. 也就是说, 除了第一项, 每一组的和都为 1/2, 所以上面的级数等于

$$1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\cdots,$$



该级数发散! 最后根据比较判别法, 调和级数发散, 因为它大于上面的发散级数. 现在我们可以轻松地知道 $\sum\limits_{n=1}^{\infty}1/n^p$ 当 $p\leqslant 1$ 时发散, 因为 $1/n^p\geqslant 1/n$, 再次应用比较判别法即可. (细节试着自行完成.)

22.4.4 绝对收敛判别法

若级数 $\sum\limits_{n=1}^{\infty}a_n$ 的各项 a_n 有的为正, 有的为负, 则会使问题变得更难了 (或更

有意思了, 这取决于你怎么看). 如果级数从某一项后都为正, 这就没问题, 可以略去前面的项, 只讨论后面的正项组成的新级数. 要知道, 级数的前面有限项不影响级数最终的敛散性. 类似地, 若级数从某一项后均为负, 可以忽略前面的有限项, 只讨论由后面的负项组成的级数. 然后, 考虑所有项均为正的级数 $\sum\limits_{n=m}^{\infty}(-a_n)$: 若该级数收敛, 则原级数也收敛; 若它发散, 则原级数也发散. 这是因为新级数与原级数相差一个负号.

若级数各项正负交错出现会怎样呢? 例如下面的例子:

$$\sum_{n=3}^{\infty}\sin(n)\left(\frac{1}{2}\right)^n, \quad \sum_{n=1}^{\infty}\frac{(-1)^n}{n^2} \text{ 和 } \sum_{n=1}^{\infty}\frac{(-1)^n}{n}.$$

第二个和第三个级数实际上是交错级数, 即各项正数和负数交错出现. 例如, 第三个级数可展开为

$$-1+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}+\cdots,$$

可以清楚地看到每隔一项为负. 上面的第一个级数不是交错的. 虽然 $\sin(n)$ 有时为正, 有时为负, 但正负项不是交错出现. 例如, $\sin(1)$、$\sin(2)$ 和 $\sin(3)$ 都是正的 (因为 1, 2 和 3 都在 0 与 π 之间), 而 $\sin(4)$、$\sin(5)$ 和 $\sin(6)$ 都是负的.

不管怎样, 我们将在 22.5.4 节专门讨论针对交错级数的判别法. 现在还有绝对收敛判别法: 若 $\sum\limits_{n=1}^{\infty}|a_n|$ 收敛, 则 $\sum\limits_{n=1}^{\infty}a_n$ 也收敛. 同样, 级数可以从 n 的任何值开始, 不必一定从 $n=1$ 开始. 现在利用该判别法讨论上面的例子. 对于第一个级数

$$\sum_{n=3}^{\infty}\sin(n)\left(\frac{1}{2}\right)^n,$$

加绝对值后为

$$\sum_{n=3}^{\infty}|\sin(n)|\left(\frac{1}{2}\right)^n.$$

注意我们只需对 $\sin(n)$ 加绝对值号, 因为因式 $(1/2)^n$ 恒正. 我们已经在 22.4.1 节用比较判别法证明了上述级数是收敛的, 则由绝对收敛判别法知原级数 (不带绝对值号的) 也收敛. 实际上, 原级数**绝对收敛**. 更多信息见 22.5.4 节.

对于第二个级数

$$\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2},$$

加绝对值后为

$$\sum_{n=1}^{\infty}\frac{1}{n^2}.$$



由 p 判别法知该级数收敛 (因为 $2>1$), 故由绝对收敛判别法知原级数绝对收敛.

对于第三个级数

$$\sum_{n=1}^{\infty}\frac{(-1)^n}{n},$$

加绝对值后为

$$\sum_{n=1}^{\infty}\frac{1}{n}.$$

由 p 判别法知该级数发散, 故不能运用绝对收敛判别法. 即不能得出原级数

$$\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$$

发散的结论. 只能说该级数不绝对收敛. 实际上, 在 22.5.4 节我们将会知道该级数是收敛的, 尽管它加上绝对值后是发散的! 在此之前, 我们还有一些其他判别法要讨论.

22.5　级数的新判别法

我们来看四个与反常积分无对应的级数收敛性判别法：比式判别法、根式判别法. 积分判别法和交错级数判别法. 在下一章讨论应用之前, 我们先来看看它们的内容.

22.5.1　比式判别法 (理论)

这是一个只能用于级数而不能用于反常积分的非常有用的判别法, 被称为比式判别法, 因为它涉及级数相邻两项的比. 提出的问题是：对于级数 $\sum\limits_{n=1}^{\infty}a_n$, 若要使它收敛, 则各项要以足够快的速度趋于 0. 解决方法是：考虑一个新的数列 b_n, 定义其为级数相邻两项之比的绝对值, 即对每个 n 令

$$b_n=\left|\frac{a_{n+1}}{a_n}\right|.$$

这是一个数列, 所以它可能收敛于某数. 结果为：若数列 $\{b_n\}$ 收敛于一个小于 1 的数, 则立即得知级数 $\sum\limits_{n=1}^{\infty}a_n$ 收敛. 实际上它绝对收敛, 即 $\sum\limits_{n=1}^{\infty}|a_n|$ 也收敛. 另一方面, 若数列 $\{b_n\}$ 收敛于一个大于 1 的数, 则级数 $\sum\limits_{n=1}^{\infty}a_n$ 发散. 若数列 $\{b_n\}$ 收敛于 1 或不收敛. 则我们对原级数得不出什么结论.

下一章将讨论比式判别法的更多例子, 现在来看能否证明该判别法. 这是一个很复杂的论证, 若不能理解, 不要着急, 可直接跳到下一节. 我们只是来试一下. 这



里假定对所有 n 有 $a_n \geqslant 0$, 这样就可以去掉绝对值号. 假设 b_n 收敛于一个小于 1 的值 L, 即

$$\text{当 } n \to \infty \text{ 时}, \quad \frac{a_{n+1}}{a_n} \to L < 1;$$

这意味着当 n 很大时, 比式 a_{n+1}/a_n 近似等于 L. 若比式就等于 L, 则级数为具有公比 L 的几何级数, 且当 $L < 1$ 时级数收敛. 但比式只是极限等于 L, 所以我们要更聪明点才行.

可令 r 等于 L 和 1 的平均值, 由于 $L < 1$, 则均值 r 介于 L 和 1 之间, 所以 r 小于 1, 即 $L < r < 1$. 然后呢? 由于比式 a_{n+1}/a_n 收敛于 L, 因而该比式会总小于 r. 也就是说, 该比式一开始可能会在某些值之间徘徊, 但最终会趋近于 L. 若比值不小于 r 就不能趋近于 L, 因为 r 大于 L. 所以, 关键点是若去除级数前面足够多项后, 总能有 a_{n+1}/a_n 小于 r.

看一下我们得到的结论: 我们从 $\sum\limits_{n=1}^{\infty} a_n$ 开始, 但去掉了前面的若干项, 得到从某数 m 开始的 $\sum\limits_{n=m}^{\infty} a_n$. 去掉前面的有限项并不影响级数的敛散性. 另一方面, 这个操作是有用的, 因为可以确定对所有 $n \geqslant m$ 有 $a_{n+1}/a_n < r$; 换一种写法即为: 对所有 $n \geqslant m$ 有 $a_{n+1} < ra_n$.

我们就要接近问题的核心了: 数列 $\{a_n\}$ 受控于公比为 r 的等比数列. 毕竟, 由 a_n 推出 a_{n+1}, 需乘以一个小于 r 的数 (因为 $a_{n+1} < ra_n$). 另一方面, 要从公比为 r 的等比数列的某项推出下一项, 也需乘以 r. 所以若等比数列从 a_m 开始, 则该数列领先于数列 $\{a_n\}$, 且保持领先. (所有这些都可用推导得出. 假设 $a_n < Ar^n$, 则两边同乘 r 可得 $ra_n < Ar^{n+1}$. 由于 $a_{n+1} < ra_n$, 我们有 $a_{n+1} < Ar^{n+1}$. 现在只需选择使得 $a_m < Ar^m$ 的 A 即可, 任何大于 a_m/r^m 的数都行.)

好的, 我们已经得到对某数 A 有 $a_n < Ar^n$, 于是

$$\sum_{n=m}^{\infty} a_n \leqslant \sum_{n=m}^{\infty} Ar^n.$$

因为 $0 \leqslant r < 1$, 右边收敛, 故由比较判别法知左边也收敛. 最后, 由绝对收敛判别法知级数 $\sum\limits_{n=m}^{\infty} a_n$ 也收敛, 虽然有些项 a_n 是负的.

真不容易, 幸运的是发散的情形不会这么麻烦. 假定比式 $|a_{n+1}/a_n|$ 收敛于一个大于 1 的数 L, 若我们删掉足够的项之后, 只需讨论 $\sum\limits_{n=m}^{\infty} |a_n|$, 其中 m 是使得 $|a_{n+1}/a_n| > 1$ 对所有 $n \geqslant m$ 均成立的足够大的数. 也就是说, 对于所有的 $n > m$ 都有 $|a_{n+1}| > |a_n|$. 项 $|a_n|$ 随着 n 的增大而增大, 所以不可能有 $\lim\limits_{n \to \infty} a_n = 0$. 现在

只需运用第 n 项判别法即可证明 $\sum\limits_{n=m}^{\infty} a_n$ 发散, 故 $\sum\limits_{n=1}^{\infty} a_n$ 也发散.

剩下的问题就是, 证明 $L=1$ 时无明确结论. 这里有一个很好的例子：考虑级数 $\sum\limits_{n=1}^{\infty} 1/n^p$, 计算相邻项的比

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{\frac{1}{(n+1)^p}}{\frac{1}{n^p}} = \frac{n^p}{(n+1)^p} = \left(\frac{n}{n+1}\right)^p.$$

我们可以去掉绝对值号, 因为各部分均为正. 当 $n\to\infty$, 显然有 $n/(n+1)\to 1$, 所以 p 次幂仍趋于 1, 即

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left(\frac{n}{n+1}\right)^p = 1^p = 1.$$

所以, 比式的极限 L 无论 p 为何值均为 1. 我们知道当 $p>1$ 时 $\sum\limits_{n=1}^{\infty} 1/n^p$ 收敛, 而 $p\leqslant 1$ 时发散. 比式极限 $L=1$ 对这两种可能都不能判定. 这个例子足以说明若 $L=1$, 则原级数可能收敛也可能发散, 只是无从判定.

22.5.2 根式判别法(理论)

根式判别法 (也叫 n 次方根判别法) 类似于比式判别法, 考虑的不是相邻项的比, 而是第 n 项绝对值的 n 次方根, 即对给定级数 $\sum\limits_{n=1}^{\infty} a_n$, 构造新数列为

$$b_n = |a_n|^{1/n}.$$

(要知道, 某量的 $1/n$ 次幂与其 n 次方根是一回事.)现在欲知数列 $\{b_n\}$ 收敛与否, 并要求极限. 若极限值小于 1, 则级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛 (事实上, 绝对收敛). 若极限值大于 1, 则级数发散. 若极限值等于 1, 则我们对原级数得不出明确结论, 需要采用其他方法进行讨论.

我们仍用下一章的一个例子来证明该结论. 若读不懂, 可直接进入下一节. 不管怎样, 其主要思想仍是通过等比数列来讨论的. 假定 $a_n=r^n$, 则 $|a_n|$ 的 n 次方根为 $|r|$, 所以当 $|r|<1$ 时级数收敛, 而当 $|r|>1$ 时级数发散. 这里, 我们并未没明确的几何级数, 但也差不多了. 我们假定



$$\text{当 } n\to\infty \text{ 时,}\quad \lim_{n\to\infty}|a_n|^{1/n} = L < 1.$$

采用与比式判别法证明相同的逻辑, 令 r 等于 L 和 1 的平均值, 最终 $|a_n|^{1/n}<r$, 即在级数中的某一点 $n=m$ 后, $|a_n|<r^n$. 故我们有

$$\sum_{n=m}^{\infty}|a_n| \leqslant \sum_{n=m}^{\infty} r^n.$$

因为 $r<1$, 所以右边级数收敛, 我们运用比较判别法可得左边级数也收敛, 所以级数 $\sum\limits_{n=1}^{\infty} a_n$ 绝对收敛.

另一方面, 假设极限值 L 大于 1, 即

$$\text{当 } n \to \infty \text{ 时,}\quad \lim_{n\to\infty} |a_n|^{1/n} = L > 1.$$

最终对足够大的 n, 总有 $|a_n|^{1/n} > 1$, 意味着 $|a_n| > 1$. 故由第 n 项判别法知级数 $\sum_{n=1}^{\infty} a_n$ 发散, 因为通项不趋于 0.

若极限值 L 为 1, 判别法仍无效, 还是用例子 $\sum_{n=1}^{\infty} 1/n^p$ 来讨论. 由你自行证明, 即

$$\lim_{n\to\infty} \left|\frac{1}{n^p}\right|^{1/n} = \lim_{n\to\infty} n^{-p/n} = 1.$$

(把它看作洛必达类型的问题进行讨论, 该类型问题参见 14.1.5 节.) 我们知道 $\sum_{n=1}^{\infty} 1/n^p$ 对某些 p 值发散, 对其他 p 值收敛. 由此可知, 根式判别法给不出任何有用的信息, 因为无论 p 为何值, 上面的极限值都为 1.

22.5.3　积分判别法 (理论)

我们在 22.4 节讨论过, 反常积分和无穷级数之间是有联系的. 积分判别法更确定了这种联系. 特别地, 对给定的级数 $\sum_{n=1}^{\infty} a_n$, 其中 a_n 为**正且递减**. 这里 “递减” 意思是对所有 n 都有 $a_{n+1} \leqslant a_n$. (更专业地说应是 “非增”, 因为这里的不等式并不严格.) 这类级数的一个例子为 $\sum_{n=1}^{\infty} 1/n^p$, 其中 $p > 0$: 各项当然为正, 且显然递减. 我们画个一般情形的图像, 如图 22-1 所示.

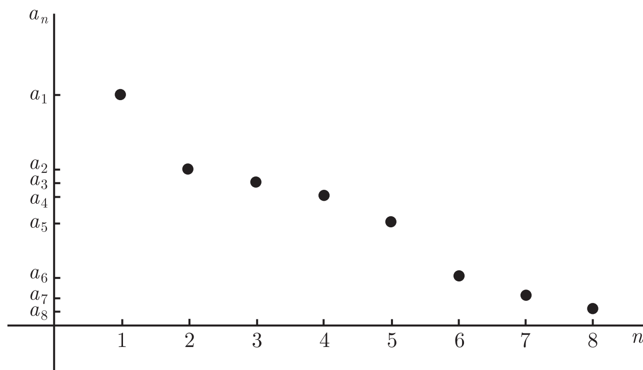


图　22-1

坐标轴用 n 和 a_n 代替 x 和 y. 数 n 上方相应点的高度是 a_n 的值, 注意所有的点都在 x 轴 (其实为 n 轴) 上方, 因为所有的 a_n 均为正. 另外, 高度在逐渐变小,

因为各项递减.

假想能找到某个递减的连续函数 f 可以把点连接起来, 如图 22-2 所示.

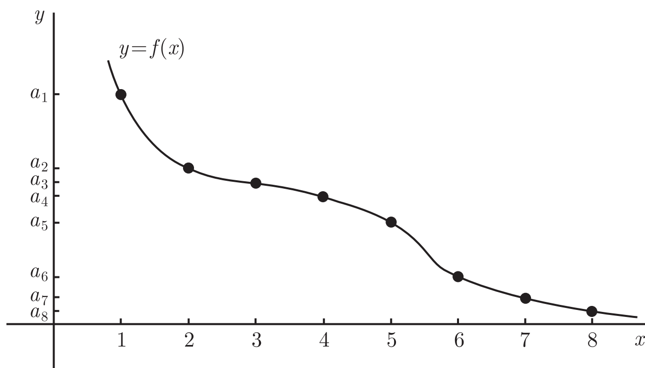


图 22-2

由于曲线 $y=f(x)$ 穿过每一个点, 所以对所有正整数 n 有 $f(n)=a_n$. 考虑积分

$$\int_1^{\infty} f(x)\mathrm{d}x.$$

若该积分收敛, 则级数 $\sum\limits_{n=1}^{\infty} a_n$ 也收敛. 为什么呢? 我们在图像上画一些线, 如图 22-3 所示.

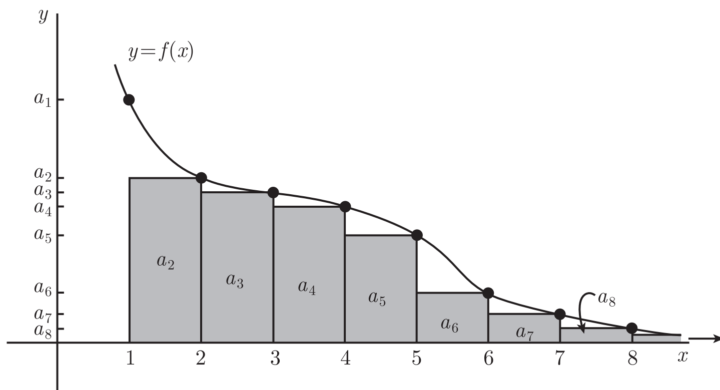


图 22-3

我们在曲线下方画了一串矩形. 每个矩形的底为 1, 高分别为 $a_2, a_3, a_4, \cdots$. (这里 a_1 没有对应的矩形.)所有矩形的总面积为 $\sum\limits_{n=2}^{\infty} a_n$ (平方单位), 由比较判别法, 该面积之和是一个有限的数, 因为

$$0 \leqslant \sum_{n=2}^{\infty} a_n \leqslant \int_1^{\infty} f(x)\mathrm{d}x < \infty.$$

所以级数 $\sum\limits_{n=2}^{\infty} a_n$ 收敛, 当然 $\sum\limits_{n=1}^{\infty} a_n$ 也收敛. (要知道, 级数的前几项不影响其收敛性!)

另一方面, 假设 $\int_1^\infty f(x)\mathrm{d}x$ 发散. 这次我们画不同的矩形, 如图 22-4 所示.

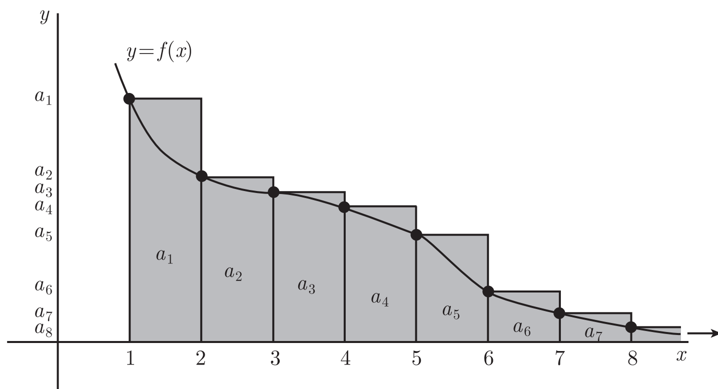


图　22-4

这次矩形延伸到曲线上方, 每个矩形的底为 1, 高分别为 $a_1, a_2, a_3, \cdots$. (这次 a_1 有对应的矩形!) 由于矩形在曲线上方, 我们有

$$\sum_{n=1}^{\infty} a_n \geqslant \int_1^{\infty} f(x)\mathrm{d}x = \infty,$$

由比较判别法可知级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

综上所述, 我们有**积分判别法**: 若 f 是使得对所有正整数 n 有 $f(n) = a_n$ 的递减正函数, 则

$$\int_1^{\infty} f(x)\mathrm{d}x \quad 和 \quad \sum_{n=1}^{\infty} a_n$$

同时收敛或同时发散. 这里, 级数还是可以从任何数开始, 不必一定从 $n = 1$ 开始, 只需要改变相应的积分下限. 下一章将讨论一些应用积分判别法的例题, 现在我们至少可以用它来证明级数的 p 判别法, 该判别法已经在 22.4.3 节见过.

为了研究 $\sum\limits_{n=1}^{\infty} 1/n^p$ 的收敛性, 首先假设 $p > 0$ 并考虑函数 $f(x) = 1/x^p$, $x > 0$. 该函数显然当 $x = n$ 时等于 $1/n^p$, 且递减. (证明递减的一个方法就是考虑导数. 在这个例子中, $f'(x) = -px^{-p-1}$, 当 $x > 0$ 时为负, 故 f 递减.) 我们可以由积分判别法得

$$\int_1^{\infty} \frac{1}{x^p}\mathrm{d}x \quad 和 \quad \sum_{n=1}^{\infty} \frac{1}{n^p}$$

同时收敛或同时发散, 到底是哪个呢? 当 $p > 1$, 根据积分的 p 判别法知积分收敛,

所以级数也收敛; 当 $0 < p \leqslant 1$, 根据积分的 p 判别法知积分发散, 所以级数也发散.

那 $p < 0$ 的时候呢? 这时不能运用积分判别法, 因为 $f(x) = 1/x^p$ 是增函数. 你知道, 若 $p < 0$, 则对于某 $q > 0$, 可令 $p = -q$. 则

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \sum_{n=1}^{\infty} \frac{1}{n^{-q}} = \sum_{n=1}^{\infty} n^q.$$

最后一个级数根据第 n 项判别法可知发散, 因为当 $n \to \infty$, $n^q \to \infty$ (不是 0). 最后, 若 $p = 0$, 级数 $\sum\limits_{n=1}^{\infty} 1/n^p$ 为 $\sum\limits_{n=1}^{\infty} 1 = 1 + 1 + 1 + \cdots$, 显然发散. 综上所述, 我们知级数 $\sum\limits_{n=1}^{\infty} 1/n^p$ 当 $p > 1$ 时收敛, 当 $p \leqslant 1$ 时发散, 这就是级数的 p 判别法!

22.5.4 交错级数判别法 (理论)

假设级数 $\sum\limits_{n=1}^{\infty} a_n$, 其中各项不确定是正还是负, 但正负交替出现. 我们在 22.4.4 节见过这样的例子, 有些情况下, 绝对收敛判别就可以解决问题, 比如当级数 $\sum\limits_{n=1}^{\infty} |a_n|$ 收敛, 则原级数收敛. 但若 $\sum\limits_{n=1}^{\infty} |a_n|$ 发散怎么办呢? 你究竟能做什么呢?

这的确是个问题, 通常没有简单的答案. 多年来, 这个困扰人的问题引起了很多的思考和讨论. 令人振奋的是, 有一个简单的判别法常被应用. 假设级数是交错的, 意味着每隔二项为正, 而每隔一项为负. 若令每个正项级数各项乘以 $(-1)^n$, 则可得到一个交错级数. (也可乘以 $(-1)^{n+1}$.) 前面我们讨论的两个级数

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \quad 和 \quad \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

都是交错级数. 我们已经知道 (见 22.4.4 节), 第一个级数绝对收敛, 故收敛. 第二个更有意思, 它**不是**绝对收敛的, 因为它的绝对值形式 $\sum\limits_{n=1}^{\infty} 1/n$ 发散. 令人惊奇的是, 原级数 $\sum\limits_{n=1}^{\infty} (-1)^n/n$ 是收敛的! 当一个级数收敛而其绝对值形式发散, 我们就说该级数*条件收敛*. 所以级数 $\sum\limits_{n=1}^{\infty} (-1)^n/n$ 条件收敛. 来看下原因.



*交错级数判别法*表明, 若级数 $\sum\limits_{n=1}^{\infty} a_n$ 是交错的, 且各项的绝对值递减趋于 0, 则级数收敛. 也就是说, a_n 正负交错, $|a_n|$ 递减, 且 $\lim\limits_{n\to\infty} |a_n| = 0$, 则级数收敛. 例如前面的级数 $\sum\limits_{n=1}^{\infty} (-1)^n/n$ 收敛, 因为它是交错级数, 各项的绝对值为递减数列 $\{1/n\}$, 且趋于 0. 在 23.7 节, 我们将总结判别法, 讨论更多关于交错级数判别法的例子.

为什么该判别法可行呢？首先我们做个可信度验证. 其中的一个条件为级数各项的极限趋于 0. 如果不是这样, 则根据第 n 项判别法可知级数发散! 所以该条件是显而易见的. 再看看剩下的条件是怎么起作用的. 考虑部分和 $\{A_N\}$, 其中 $A_N=\sum\limits_{n=1}^{N}a_n$. 由于 a_n 不停地在正负之间交错, 部分和 A_N 则来回游移. 回想持扩音器的家伙告诉你来回走：每一秒, 他告诉你向前走, 而下一秒他告诉你向后走. 你可能向前走迈右脚, 而向后走会迈左脚. 另一方面, 步长 (即 $|a_n|$) 变得越来越小且趋于 0, 所以你发现自己在用越来越小的步长来回走动. 这意味着, 你左脚和右脚正在向一起靠近. 每次迈出左脚, 就会比原来的位置远一点; 每次迈出右脚, 就会回来一点. 在极限情况下, 你的两只脚并在了同一点上, 所以级数收敛!

假定 $a_1,a_3,a_5,\cdots$ 都为正, $a_2,a_4,a_6,\cdots$ 都为负, 则我们可以用数学方式表述上述过程. 现在考虑奇部分和 $A_1,A_3,A_5,\cdots$, 这是你的右脚不断走的位置. 我要求其为递减数列. 实际上, $A_1=a_1$, 而 $A_3=a_1+a_2+a_3$, 也可写为 $A_1+a_2+a_3$. 现在 a_2 是负的, a_3 是正的, 且由对步长递减的假设知 $|a_2|\geqslant|a_3|$. 这意味着 $a_2+a_3\leqslant 0$, 即 $A_3=A_1+a_2+a_3\leqslant A_1$. 现在对 A_5 重复该论证, 看会发生什么. 你知道, A_5 是前五项 a_n 之和, 而 A_3 是前三项之和, 故可以写为 $A_5=A_3+a_4+a_5$. (若你在三步之后知道自己在哪儿, 即 A_3, 则只需走下两个带符号的步 a_4 和 a_5 来看一下五步之后在哪儿, 即 A_5.) 不管怎样, $a_4+a_5\leqslant 0$, 因为 a_4 为负, a_5 为正, 且 $|a_4|\geqslant|a_5|$. 这意味着 $A_5\leqslant A_3$. 若继续该过程, 你会发现

$$A_1\geqslant A_3\geqslant A_5\geqslant A_7\geqslant\cdots,$$

所以你的右脚实际上随着时间的流逝一直往回走.

你可以对偶部分和 $A_2,A_4,A_6,\cdots$ 重复相同的论证 (但方向相反). 做一下看看能否得出

$$A_2\leqslant A_4\leqslant A_6\leqslant A_8\leqslant\cdots,$$

故你的左脚随时间变化不断向前走. 关键是：数列 $A_1,A_3,A_5,\cdots$ 递减, 所以它要么趋于 $-\infty$, 要么收敛于某有限值. 不过它不会趋于 $-\infty$, 因为所有项都大于 A_2. (为什么呢？) 类似地, 偶数列 $A_2,A_4,A_6,\cdots$ 递增, 故其要么趋于 ∞, 要么收敛. 它不会趋于 ∞, 因为所有项都小于 A_1. (为什么？) 所以奇部分和数列与偶部分和数列都收敛. 由于两者之差 $|a_n|$ 越来越小, 故两者的极限一定相同! 即, 奇部分和数列递减至偶部分和数列增长的值, 即你的两只脚靠得越来越近直到任意接近. 这就证明了部分和数列 $\{A_N\}$ 收敛, 而这意味着原级数 $\sum\limits_{n=1}^{\infty}a_n$ 也收敛.

所以交错级数判别法可行. 重要的是, 只有当你验证了给定的级数非绝对收敛时才需要用它. 下一章我们将用更多例题来讨论它的用法.

第 23 章　求解级数问题

问题：对于级数 $\sum_{n=1}^{\infty} a_n$, 确定它收敛与否. 若该级数收敛, 你可能想知道它的值 (即收敛于哪个值). 要想求得一个具有漂亮表达式的级数值, 这个级数就必须很特殊. 当然, 级数不必与上面一样从 $n=1$ 开始, 它可以从 $n=0$ 或 n 的其他值开始.

本章主要围绕这样的问题展开讨论. 下面是关于级数的讨论布图.

(1) **是几何级数吗？** 如果级数只包含 2^n 或 e^{3n} 这样的指数, 那么它可能是几何级数, 或是多个几何级数之和. 这种情形的讨论见 23.1 节.

(2) **级数各项趋于 0 吗？** 如果不是几何级数, 尝试用**第 n 项判别法**. 检验一下各项是否趋于 0, 如果不是, 则根据第 n 项判别法知级数发散. 详情见 23.2 节.

(3) **级数中有负项吗？**如果有, 你可能要用**绝对收敛判别法**或**交错级数判别法**. 更多信息见 23.7 节.

(4) **级数中有阶乘吗？**如果有, 用**比式判别法**. 该判别法同样适用于级数中包含指数而非几何级数的情形. 详情见 23.3 节.

(5) **有底和指数都包含 n 的指数吗？**如果有, 试着用**根式判别法**. 一般地, 如果容易对项 a_n 取 n 次方根, 就可以用根式判别法. 详情见 23.4 节.

(6) **项里面有因子 $1/n$ 或对数吗？**在这种情况下, **积分判别法**可能是你想用的. 我们将在 23.5 节讨论.

(7) **上面的所有判别法都不能用吗？**你可能需要用**比较判别法**或 **p 判别法**与**极限比较判别法**联合使用, 并重温第 21 章关于函数所有表现的讨论. 我们将在 23.6 节应用这些判别法.

以上讨论计划将引导你在各种不同的级数中穿梭. 上述这些其实并不完美, 仍不时会有陷阱出现, 希望这些情况能少出现一点. 我的建议是掌握所有这些资料, 然后在你平时的学习中时刻提防不常见的陷阱. 现在让我们来看一下具体细节.

23.1　求几何级数的值

如果级数只包含 2^n 或 e^{3n} 这样的指数, 那么它可能是一个或多个几何级数之和. 如上一章所述, 几何级数很简单, 可以直接求它的值 (如果它收敛的话). 几何级数的一般式是 $\sum_{n=m}^{\infty} ar^n$, 其中 r 是公比. 在 22.2.1 节, 我们讨论了如何求该级数的值, 我推荐用文字而不是数学语言来学习这个方法：

$$\text{若 } -1 < \text{公比} < 1, \text{无穷几何级数的和} = \frac{\text{首项}}{1-\text{公比}}.$$



如果公比不是介于 −1 和 1 之间, 则级数发散.

我们来具体应用, 假定要求解

$$\sum_{n=5}^{\infty} \frac{4}{3^n}.$$

这是一个几何级数, 因为我们有

$$\frac{4}{3^n} = 4\left(\frac{1}{3}\right)^n.$$

由此, 我们可知公比是 1/3, 它介于 −1 到 1 之间, 故该级数收敛. 你会问：收敛于何处？首项在 $n = 5$ 时为 $4/3^5$. 所以

$$\sum_{n=5}^{\infty} \frac{4}{3^n} = \sum_{n=5}^{\infty} 4\left(\frac{1}{3}\right)^n = \frac{4/3^5}{1-1/3},$$

结果为 2/81.



这是个极具欺骗性的例子：

$$\sum_{n=2}^{\infty} \frac{2^{2n}-(-7)^n}{11^n}.$$

它不是几何级数, 但可以分成两个几何级数之差：

$$\sum_{n=2}^{\infty} \frac{2^{2n}-(-7)^n}{11^n} = \sum_{n=2}^{\infty} \frac{2^{2n}}{11^n} - \sum_{n=2}^{\infty} \frac{(-7)^n}{11^n}.$$

为什么分开后两个都是几何级数呢？在第一个级数中, 你可以用 4^n 替换 2^{2n}, 然后将 $4^n/11^n$ 表示成 $(4/11)^n$. 最后这一步变换同样可以用在第二个级数中, 我们有

$$\sum_{n=2}^{\infty} \frac{2^{2n}-(-7)^n}{11^n} = \sum_{n=2}^{\infty} \left(\frac{4}{11}\right)^n - \sum_{n=2}^{\infty} \left(\frac{-7}{11}\right)^n.$$

由于这两个级数的公比分别为 4/11 和 −7/11, 都介于 −1 和 1 之间, 因此都收敛, 故我们可以写成上式. $n = 2$ 为第一项, 所以首项分别为 $(4/11)^2$ 和 $(-7/11)^2$. 因此计算出来为

$$\frac{(4/11)^2}{1-(4/11)} - \frac{(-7/11)^2}{1-(-7/11)},$$

化简后为 −5/126.



如果把问题稍微变一下会怎样呢？考虑

$$\sum_{n=2}^{\infty}\frac{2^{2n}-(-13)^n}{11^n}.$$

这次还是把和拆成两组, 写为

$$\sum_{n=2}^{\infty}\left(\frac{4}{11}\right)^n-\sum_{n=2}^{\infty}\left(\frac{-13}{11}\right)^n.$$

不必求出第一个级数的值, 只需要知道它收敛. 第二个级数由于公比 $-13/11$ 不是介于 -1 和 1 之间而发散. 收敛级数和发散级数之和一定发散!

可见, 几何级数很好计算. 如果给定的不是几何级数, 按照下面的顺序, 从第 n 项判别法开始进行判断.

23.2 应用第 n 项判别法

无论什么时候都要首先考虑第 n 项判别法! 其内容为:

> 若 $\lim\limits_{n\to\infty}a_n\neq 0$ 或极限不存在, 则级数 $\sum\limits_{n=1}^{\infty}a_n$ 发散.

若级数通项不趋于 0, 则级数定发散. 若通项趋于 0, 则级数可能发散也可能收敛: 需要进一步判断. **该判别法不能用于级数收敛性的判定**. 总之, 只需快速检验一下通项是否趋于 0, 避免在其他判别法上浪费时间. 例如, 考察级数



$$\sum_{n=1}^{\infty}\frac{n^2-3n+7}{4n^2+2n+1},$$

不需要考虑其他任何判别法, 只要注意

$$\lim_{n\to\infty}\frac{n^2-3n+7}{4n^2+2n+1}=\frac{1}{4},$$

即该级数通项不趋于 0, 由第 n 项判别法可知原级数发散.

如果级数通项趋于 0, 则需要尝试用其他判别法来判别. 在进行判定之前, 一定要快速看下级数是否有负项. 这种情况一般发生在有些项包含负号、因子 $(-1)^n$ 或三角函数 (尤其是 $\sin(n)$ 或 $\cos(n)$) 时. 出现负项的情形参见 23.7 节. 若各项均为正, 用下面的判别法进行判定.



23.3 应用比式判别法

若级数中包含阶乘, 用比式判别法. 记住, 阶乘包含感叹号, 如 $n!$ 或 $(2n+5)!$.

对于含有指数的级数, 如 2^n 或 $(-5)^{3n}$, 比式判别法同样适用. 根据 22.5.1 节, 该判别法总结为：

> 若 $L = \lim_{n\to\infty}\left|\dfrac{a_{n+1}}{a_n}\right|$, 则 $n = \sum_{n=1}^{\infty} a_n$ 在 $L < 1$ 时绝对收敛, 在 $L > 1$ 时发散; 但当 $L = 1$ 或极限不存在时, 比式判别法无效.

可按照下面的步骤使用比式判别法进行判别：

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{\text{第 } n \text{ 项中用 } n+1 \text{ 代换 } n}{\text{第 } n \text{ 项}}\right|$$

因为可能会是分式比分式的形式, 所以这里要用稍长的分数线. 级数的第 n 项是 a_n, 把 n 换成 $(n+1)$ 就得到 a_{n+1}. 现在求上面的极限. 假设我们已经完成这一步并求得极限值 L, 则有三种可能：

(1) 若 $L < 1$, 则原级数 $\sum_{n=1}^{\infty} a_n$ 收敛, 实际上是绝对收敛;

(2) 若 $L > 1$, 则原级数发散;

(3) 若 $L = 1$ 或极限不存在, 则比式判别法无效, 尝试用其他方法.

来看一些例子. 首先考虑

$$\sum_{n=1}^{\infty}\frac{n^{1000}}{2^n}.$$

由于分子是多项式, 所以该级数不是几何级数, 又因为指数增长速度快于多项式 (见 21.3.3 节), 可知第 n 项的极限为 0, 即

$$\lim_{n\to\infty}\frac{n^{1000}}{2^n} = 0,$$

则我们不能用第 n 项判别法. 因为该级数包含指数, 我们试一下比式判别法. 根据标准步骤, 有

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{\frac{(n+1)^{1000}}{2^{n+1}}}{\frac{n^{1000}}{2^n}}\right|.$$

注意分母就是第 n 项, 我们直接将其从原级数中挪过来的; 分子除了将 n 换为 $n+1$ 外, 跟分母一样. 现在通过将上述表达式颠倒相乘, 分组同类项, 得到

$$\lim_{n\to\infty}\left|\frac{(n+1)^{1000}}{n^{1000}}\frac{2^n}{2^{n+1}}\right| = \lim_{n\to\infty}\left(\frac{n+1}{n}\right)^{1000}\frac{1}{2} = 1^{1000}\times\frac{1}{2} = \frac{1}{2}.$$

注意我们去掉了绝对值号 (每项都为正), 把 1000 次方的项也写在一起, 同时运用了事实 $\lim_{n\to\infty}(n+1)/n = 1$. 总之, 上面的极限为 1/2, 它小于 1, 所以由比式判别法可知原级数收敛. 解题完毕.



考虑

$$\sum_{n=2}^{\infty}\frac{3^n}{n\ln(n)}.$$

能够得到, 当 $n\to\infty$ 时, 级数通项趋于 ∞, 所以由第 n 项判别法可知该级数发散. 假设你只考虑了比式判别法, 同样能得到

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=\lim_{n\to\infty}\left|\frac{\frac{3^{n+1}}{(n+1)\ln(n+1)}}{\frac{3n}{n\ln(n)}}\right|=\lim_{n\to\infty}\frac{3^{n+1}}{3^n}\frac{n}{n+1}\frac{\ln(n)}{\ln(n+1)}=3.$$

我们用到了 $\lim\limits_{n\to\infty}n/(n+1)=1$ 和 $\lim\limits_{n\to\infty}\ln(n)/\ln(n+1)=1$, 前者好求, 而后者就不容易了. 对后一个极限, 可用洛必达法则验证是否极限为 1. 总之, 上述比值的极限是 3, 因 $3>1$, 原级数发散. 所以, 虽然我们没用第 n 项判别法, 比式判别法也能判别.



当遇到含阶乘的级数时, 比式判别法是相当有用的. 记住, $n!$ 是从 1 到 n 的自然数之积:

$$n!=1\times2\times3\times\cdots\times(n-1)\times n.$$

当用比式判别法对含阶乘的级数判别时, 经常会考虑到比式

$$\frac{n!}{(n+1)!}.$$

化简该式的唯一可行方法就是将阶乘展开并做相消, 即

$$\frac{n!}{(n+1)!}=\frac{1\times2\times\cdots\times(n-1)\times n}{1\times2\times\cdots\times(n-1)\times n\times(n+1)}=\frac{1}{n+1}.$$

该方法还可以, 不过当遇到类似 $(2n)!$ 的式子可能就有麻烦了. $(2n)!$ 与 $2\times n!$ 不同 —— 这是个常犯的错误. 考虑比式

$$\frac{(2(n+1))!}{(2n)!}=\frac{(2n+2)!}{(2n)!}.$$

分子是前 $2n+2$ 个数之积, 而分母只是前 $2n$ 个数之积. 故比式为

$$\frac{1\times2\times\cdots\times(2n-1)\times(2n)\times(2n+1)\times(2n+2)}{1\times2\times\cdots\times(2n-1)\times(2n)}=(2n+1)(2n+2).$$



这类计算经常有, 例如级数

$$\sum_{n=1}^{\infty}\frac{(2n)!}{(n!)^2}.$$

该级数收敛还是发散呢？通项是否趋于 0 并不清楚, 但级数中含有阶乘, 所以我们直接考虑用比式判别法：

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{\frac{(2(n+1))!}{((n+1)!)^2}}{\frac{(2n)!}{(n!)^2}}\right| = \lim_{n\to\infty}\frac{(2n+2)!}{(2n)!}\left(\frac{n!}{(n+1)!}\right)^2.$$

注意我们对比式和幂次进行了变形, 上述结果可化简为

$$\lim_{n\to\infty}(2n+2)(2n+1)\left(\frac{1}{n+1}\right)^2 = \lim_{n\to\infty}\frac{4n^2+6n+2}{n^2+2n+1} = 4.$$

所以极限值大于 1, 级数发散. 为了说明该级数的敏感性, 我们对其做个小小的修改, 在其通项的分母上加一个因子 5^n, 为

$$\sum_{n=1}^{\infty}\frac{(2n)!}{5^n(n!)^2}.$$

现在试着计算比值, 将额外的一个因子 1/5 提出来, 就可求得极限为 4/5, 小于 1, 所以修改后的级数收敛.

考虑级数

$$\sum_{n=1}^{\infty}\frac{n!}{(n+3)^n}.$$

它含有阶乘, 故我们用比式判别法, 有

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{\frac{(n+1)!}{((n+1)+3)^{n+1}}}{\frac{n!}{(n+3)^n}}\right| = \lim_{n\to\infty}\frac{(n+1)!}{n!}\frac{(n+3)^n}{(n+4)^{n+1}}.$$

可将上式中的 $(n+1)!/n!$ 化简为 $(n+1)$, 因此, 上式结果可化为

$$\lim_{n\to\infty}(n+1)\frac{(n+3)^n}{(n+4)^{n+1}}.$$

现在怎么做呢？看起来似乎很难. 何不把分母变为 $(n+4)\times(n+4)^n$, 以使其与分子的幂次一样呢？然后, 我们就可以把各部分组合成：

$$\lim_{n\to\infty}(n+1)\frac{(n+3)^n}{(n+4)^{n+1}} = \lim_{n\to\infty}\frac{n+1}{n+4}\frac{(n+3)^n}{(n+4)^n} = \lim_{n\to\infty}\frac{n+1}{n+4}\left(\frac{n+3}{n+4}\right)^n.$$

有些明朗了. 第一个因子 $(n+1)/(n+4)$ 显然在 $n\to\infty$ 时趋于 1, 但第二个似乎有点难. 计算它的一个方法就是将 n 换成 x, 考虑极限

$$\lim_{x\to\infty}\left(\frac{x+3}{x+4}\right)^x.$$

根据洛必达法则类型 C (参见 14.1.5 节), 求对数 (经过某些代数运算后) 的极限：

$$\lim_{x\to\infty}\ln\left(\frac{x+3}{x+4}\right)^x = \lim_{x\to\infty}x\ln\left(\frac{x+3}{x+4}\right) = \lim_{x\to\infty}\frac{\ln\left(\frac{x+3}{x+4}\right)}{1/x}.$$

分子当 $x\to\infty$ 时趋于 0, 因为 $(x+3)/(x+4)\to 1$, 且 $\ln(1)=0$. 分母也趋于 0, 可用洛必达法则证明

$$\lim_{x\to\infty} \ln\left(\frac{x+3}{x+4}\right)^x = -1.$$

这部分工作留给你自己完成. 取幂并将 x 换回 n, 我们就得到

$$\lim_{n\to\infty}\left(\frac{n+3}{n+4}\right)^n = \mathrm{e}^{-1}.$$

现在, 我们需要的每一部分都已得到, 上面比式的极限值为

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\frac{n+1}{n+4}\left(\frac{n+3}{n+4}\right)^n = 1\times\mathrm{e}^{-1} = \frac{1}{\mathrm{e}}.$$

由于该极限值小于 1, 故原级数收敛.



级数

$$\sum_{n=2}^{\infty}\frac{1}{n\ln(n)}$$

的敛散性如何呢? 显然, 当 $n\to\infty$ 时通项趋于 0. 我们试一下比式判别法:

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{\frac{1}{(n+1)\ln(n+1)}}{\frac{1}{n\ln(n)}}\right| = \lim_{n\to\infty}\frac{n}{n+1}\frac{\ln(n)}{\ln(n+1)} = 1.$$

(对于含对数的比式的极限, 仍要用到洛必达法则.) 我们已经得到比式的极限为 1, 这意味着什么? 这就意味着比式判别法不能给出任何有用的信息. 除了知道比式判别法无效外, 与刚看到级数时状况相比, 我们没有得到更多关于级数的信息. 我们需要试一下其他的方法. 事实上, 积分判别法是判定这个级数的最好方法, 我们将在 23.5 节进行讨论.

23.4　应用根式判别法

当级数通项的指数为特殊的关于 n 的函数时, 用根式判别法. 当通项具有形式 A^B, 其中 A 和 B 都为关于 n 的函数时, 根式判别法尤其有用. 根据 22.5.2 节, 该判别法的新内容为:

> 若 $L=\lim\limits_{n\to\infty}|a_n|^{1/n}$, 则 $n=\sum\limits_{n=1}^{\infty}a_n$ 在 $L<1$ 时绝对收敛, 在 $L>1$ 时发散; 但当 $L=1$ 或极限不存在时, 根式判别法无效.



使用根式判别法, 一般先讨论表达式

$$\lim_{n\to\infty}|a_n|^{1/n},$$

然后将 a_n 代换为所研究级数的通项, 求极限 (如果存在的话), 并称之为 L. 则与比式判别法一样, 有三种可能, 幸好结论也是一样的:

(1) 若 $L<1$, 则原级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, 实际上是绝对收敛;

(2) 若 $L>1$, 则原级数发散;

(3) 若 $L=1$ 或极限不存在, 则根式判别法无效, 尝试用其他方法.

例如, 考虑级数

$$\sum_{n=1}^{\infty}\left(1-\frac{2}{n}\right)^{n^2}.$$

由于通项的指数包含 n 的幂次, 因而该级数能用根式判别法来判别. 应用根式判别法, 有

$$\lim_{n\to\infty}|a_n|^{1/n}=\lim_{n\to\infty}\left|\left(1-\frac{2}{n}\right)^{n^2}\right|^{1/n}=\lim_{n\to\infty}\left(1-\frac{2}{n}\right)^{n^2\times\frac{1}{n}}=\lim_{n\to\infty}\left(1-\frac{2}{n}\right)^{n}$$
$$=\mathrm{e}^{-2}<1.$$

注意我们把绝对值号去掉了, 因为该式为正, 并应用了 22.1.2 节最后的那个重要极限 (把 k 换成 -2). 所以极限值为 e^{-2}, 显然小于 1; 由根式判别法知原级数收敛.

23.5 应用积分判别法

当级数同时含有 $1/n$ 和 $\ln(n)$ 时, 可用积分判别法. 由 22.5.3 节知, 若 N 为任意正整数, 则我们可以说:

若对连续递减函数 f 有 $a_n=f(n)$, 则 $\sum\limits_{n=N}^{\infty} a_n$ 与 $\int_N^{\infty} f(x)\mathrm{d}x$ 同时收敛或同时发散.

下面是积分判别法的实际应用步骤.

- 将 n 换为 x, 将 $\sum\limits_{n=1}^{\infty}$ 换成 $\int_1^{\infty}$, 并在后面加 $\mathrm{d}x$. 当然, 若级数从 $n=2$ 开始, 则用 $\int_2^{\infty}$ 代换.
- 检验被积函数是否递减, 这可以通过验证导数是否为负或直接检查被积函数获知.
- 现在讨论第一步中的反常积分. 用积分讨论级数的一个主要好处是可以对积分做换元 (或变量替换, 如果你喜欢的话). 本书中最常见的换元是 $t=\ln(x)$.
- 若反常积分收敛, 则级数也收敛; 若积分发散, 则级数也发散.

例如, 考虑

$$\sum_{n=2}^{\infty}\frac{1}{n\ln(n)}.$$

事实上我们已经讨论过该级数了, 在 23.3 节曾试图用比式判别法进行敛散性判定, 但没有成功. 现在我们换用积分判别法进行判定, 因为它包含了 $1/n$ 和 $\ln(n)$. 将变量 n 换为 x, 和式换为积分, 可得

$$\int_2^\infty \frac{1}{x\ln(x)}\mathrm{d}x.$$

被积函数 $1/(x\ln(x))$ 关于 x 递减, 这可以通过证明其导数为负得到; 或更直接地, 观察 x 和 $\ln(x)$ 均关于 x 递增, 则它们的乘积也递增, 所以倒数 $1/(x\ln(x))$ 递减. 不管怎样, 我们已经在第 21 章讨论过这个积分, 这里给出大致过程: 做换元 $t=\ln(x)$, 则 $\mathrm{d}t=1/x\mathrm{d}x$, 积分变为

$$\int_{\ln(2)}^\infty \frac{1}{t}\mathrm{d}t,$$

由积分 p 判别法可知该积分发散. 由于积分发散, 则原级数也发散 (积分判别法).



另一方面, 我们对级数做个小的变动：考虑

$$\sum_{n=2}^\infty \frac{1}{n(\ln(n))^2}.$$

同样包含因子 $1/n$ 和对数, 所以尝试用积分判别法. 将 n 换为 x, 并将级数转换为积分可得

$$\int_2^\infty \frac{1}{x(\ln(x))^2}\mathrm{d}x.$$

确保被积函数关于 x 递减. 做换元 $t=\ln(x)$, 这次积分变为

$$\int_{\ln(2)}^\infty \frac{1}{t^2}\mathrm{d}t,$$

根据 p 判别法, 该积分收敛. 这一次级数收敛 (积分判别法). 将这个例子与前面例子一起来看, 可以发现级数收敛的整个过程是多么地微妙. 我们知道随着 n 的增大, $\ln(n)$ 与 n 的任意正数次幂相比是很小的, 但上面的例子共同说明了对数会带来很大的不同. $\sum\limits_{n=2}^{\infty} 1/(n\ln(n))$ 分母上 $\ln(n)$ 的幂次增加一个很小的量都会将一个发散级数变为一个收敛级数. (在 21.3.4 节, 我们见过一个类似的例子.)

23.6　应用比较判别法、极限比较判别法和 p 判别法

当其他的判别法不能使用时, 对正项级数应用这些判别法. 你一定是想最先应用第 n 项判别法, 然后对包含阶乘的级数采用比式判别法, 对包含底和指数都为 n 的函数的项的级数采用根式判别法, 或对包含因子 $1/n$ 和对数的级数采用积分判

别法. 还剩下什么呢? 基本上与积分的工具一样: 比较判别法、极限比较判别法、p 判别法, 以及对常见函数在 ∞ 和 0 附近的表现的理解. 非常有必要在学习本节前复习第 21 章, 因为方法几乎是相同的. 不管怎样, 这里会再次讨论那些判别法. (为了便于对比, 在这里, 比较判别法和极限比较判别法中的 a_n 都假定为非负.)

(1) **比较判别法的发散情形**: 若认为 $\sum\limits_{n=1}^{\infty} a_n$ 发散, 则找一个同样发散的较小的级数, 即找一个使得对所有 n 都有 $a_n \geqslant b_n$ 的正项数列 $\{b_n\}$, 使得级数 $\sum\limits_{n=1}^{\infty} b_n$ 发散. 则

$$\sum_{n=1}^{\infty} a_n \geqslant \sum_{n=1}^{\infty} b_n = \infty,$$

所以级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

(2) **比较判别法的收敛情形**: 若认为 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, 则找一个同样收敛的较大的级数; 即找一个使得对所有 n 都有 $a_n \leqslant b_n$ 的数列 $\{b_n\}$, 使得级数 $\sum\limits_{n=1}^{\infty} b_n$ 收敛. 则

$$\sum_{n=1}^{\infty} a_n \leqslant \sum_{n=1}^{\infty} b_n < \infty,$$

所以级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

(3) **极限比较判别法**: 找一个当 $n \to \infty$ 时$a_n \sim b_n$ 的简单级数 $\sum\limits_{n=1}^{\infty} b_n$. 则若 $\sum\limits_{n=1}^{\infty} b_n$ 收敛, $\sum\limits_{n=1}^{\infty} a_n$ 也收敛. 另一方面, 若 $\sum\limits_{n=1}^{\infty} b_n$ 发散, 则 $\sum\limits_{n=1}^{\infty} a_n$ 也发散. (要知道 "当 $n \to \infty$ 时 $a_n \sim b_n$" 与 "$\lim\limits_{n\to\infty} a_n/b_n = 1$" 意思一样.)

(4) $\boldsymbol{p}$ **判别法**: 若 $a \geqslant 1$, 级数

$$\sum_{n=a}^{\infty} \frac{1}{n^p} \begin{cases} \text{收敛}, & \text{如果 } p > 1; \\ \text{发散}, & \text{如果 } p \leqslant 1. \end{cases}$$

这与积分的 p 判别法中 $\int^{\infty}$ 情形一样.



现在来看一些例子. 在下面的每个例子中, 你都可以用积分代换和式得到一个反常积分 (有瑕点 ∞) 来代换级数. 反常积分问题的解就是相应级数的解. 对每种情形, 应该试着写下对等的反常积分的问题和解. 返回第 21 章, 试着将每个瑕点为 ∞ 的反常积分转换成级数, 也是一个好办法. 它们几乎都可以用上述判别法求解. (解中包含变量变换 $t = \ln(x)$ 的问题是个例外. 对于这些问题, 为了求解相应的级数问题, 你需要用积分判别法.) 考虑级数



$$\sum_{n=1}^{\infty} \frac{2n^2+3n+7}{n^4+2n^3+1}.$$

为了检验这个说法, 注意每个多项式的最高次项起决定作用, 由于 n 变得越来越大 (详见 21.3.1 节), 我们有

$$\text{当 } n \to \infty \text{ 时}, \quad \frac{2n^2+3n+7}{n^4+2n^3+1} \sim \frac{2n^2}{n^4} = \frac{2}{n^2}.$$

由 p 判别法知 $\sum\limits_{n=1}^{\infty} 2/n^2$ 收敛 (常数 2 不相关); 故由极限比较判别法知原级数也收敛.

考虑几乎一样的例子:

$$\sum_{n=0}^{\infty} \frac{2n^2+3n+7}{n^4+2n^3+1},$$

这需要采用一点小小的技巧. 该级数与上个级数的唯一区别是, 和式从 $n=0$ 开始. 若采用与前面级数相同的讨论方法, 就会发现需将上述级数与 $\sum\limits_{n=0}^{\infty} 2/n^2$ 进行比较. 后者显然没有定义好, 因为它的第一项看似为 $2/0$, 显然没有意义. 你可以通过如下两种方法之一来避免这样的问题: 可以改变首项 $n=0$, 如换成 $n=1$, 这样并不改变原级数的敛散性; 或者, 将首项从和式中提出来. 事实上, 当 $n=0$, $(2n^2+3n+7)/(n^4+2n^3+2)$ 为 7, 所以

$$\sum_{n=0}^{\infty} \frac{2n^2+3n+7}{n^4+2n^3+1} = 7 + \sum_{n=1}^{\infty} \frac{2n^2+3n+7}{n^4+2n^3+1}.$$

右边的级数收敛, 故左边的级数也收敛. 加上有限数 7 仍然收敛. 通常, 若和式从 $n=0$ 开始, 且你想应用极限比较判别法, 就可将首项提出来, 这样就可以考虑从 $n=1$ 开始的级数了.

现在来看

$$\sum_{n=1}^{\infty} \frac{\sqrt[3]{27n^6+9n^2+4}}{n^3+9n^2+4}.$$

根据我们关于较高次幂起决定作用的标准观点, 有

$$\text{当 } n \to \infty \text{ 时}, \quad \frac{\sqrt[3]{27n^6+9n^2+4}}{n^3+9n^2+4} \sim \frac{\sqrt[3]{27n^6}}{n^3} = \frac{3n^2}{n^3} = \frac{3}{n}.$$

由 p 判别法知 $\sum\limits_{n=1}^{\infty} 3/n$ 发散, 故由极限比较判别法知原级数也发散.

那

$$\sum_{n=1}^{\infty} 2^{-n} n^{1000}$$

呢？在 23.3 节，我们运用了比式判别法来解这个问题. (事实上我们将 $2^{-n}n^{1000}$ 写成了 $n^{1000}/2^n$，不过它们当然是一样的!) 现在我们用比较判别法来求解这个问题. 用这种方法解题，需要用到指数增长较快的观点. 用 21.3.3 节描述的方法，我们有

$$2^{-n} \leqslant \frac{C}{n^{1002}},$$

这里选择指数为 1002，因为它比问题中的指数 1000 大 2. 现在我们有

$$\sum_{n=1}^{\infty} 2^{-n}n^{1000} \leqslant \sum_{n=1}^{\infty} \frac{C}{n^{1002}}n^{1000} = C\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

最后的级数由 p 判别法可知收敛，故由比较判别法知原级数也收敛.

现在考虑

$$\sum_{n=2}^{\infty} \frac{\ln(n)}{n^{1.001}}.$$

这恰恰是 21.3.4 节中例子的级数形式. 事实上，你可以用积分判别法将该级数问题转化为反常积分问题，因为被积函数是递减的，但关键点是什么？我们可以直接解该题. 与在反常积分中的方法一样，我们采用 $\ln(n) \leqslant Cn^{0.0005}$，这里巧妙地选择如此小的指数 0.0005 以使得原指数 (原级数中的指数)1.001 减去它后仍大于 1. 故我们有

$$\sum_{n=1}^{\infty} \frac{\ln(n)}{n^{1.001}} \leqslant \sum_{n=1}^{\infty} \frac{Cn^{0.0005}}{n^{1.001}} = C\sum_{n=1}^{\infty} \frac{1}{n^{1.0005}} < \infty,$$

最后的级数由 p 判别法可知收敛，故根据比较判别法知原级数收敛.

级数

$$\sum_{n=1}^{\infty} \frac{|\sin(n)|}{n^2}$$

相当容易求解. 要知道 $|\sin(n)| \leqslant 1$，我们知

$$\sum_{n=1}^{\infty} \frac{|\sin(n)|}{n^2} \leqslant \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

故由比较判别法知级数收敛.

现在考虑级数

$$\sum_{n=1}^{\infty} \sin\left(\frac{1}{n}\right).$$

该级数似乎有些项为负，不过那只不过是表面现象. 事实上，当 n 从 1 开始随着正整数的不断增大，数 $1/n$ 从 1 开始不断减小并趋于 0. 即，$1/n$ 总是介于 0 和 1 之间. 由于 $\sin(x)$ 在 0 和 1 之间恒正，则级数所有项均为正. 我们还没解决该问题，下面该怎么做呢？ 在 21.4.2 节，当 $x \to 0$ 时 $\sin(x) \sim x$. 用 $1/n$ 替换 x，则当

$1/n \to 0$ 时 $\sin(1/n) \sim 1/n$. 等一下, 当 $1/n \to 0$ 必须使 $n \to \infty$. 即, 我们已经证得当 $n \to \infty$ 时 $\sin(1/n) \sim 1/n$, 这正是我们需要的! 由于级数 $\sum\limits_{n=1}^{\infty} 1/n$ 发散, 由极限比较判别法知原级数也发散. (将这个例题与 21.4.2 节中最后一个例题比较一下.)

另一方面, 级数

$$\sum_{n=1}^{\infty} \sin^2\left(\frac{1}{n}\right)$$

收敛, 因为当 $n \to \infty$ 时 $\sin^2(1/n) \sim 1/n^2$, 具体过程请自行完成.

最后来看一个很让人头疼的级数

$$\sum_{n=2}^{\infty} \cos^2(n) \tan\left(\frac{(n^2+4n-3)\ln(n)}{\sqrt{n^7+2n^4+3n}}\right).$$

如何讨论呢? 分开考虑该级数. 当 $n \to \infty$, 因式 (n^2+4n-3) 渐近等价于 n^2, 且因式 $\sqrt{n^7+2n^4+3n}$ 渐近等价于 $\sqrt{n^7}$, $\sqrt{n^7}$ 就是 $n^{7/2}$. 所以可以说

$$\text{当 } n \to \infty \text{ 时,} \quad \frac{(n^2+4n-3)\ln(n)}{\sqrt{n^7+2n^4+3n}} \sim \frac{n^2\ln(n)}{n^{7/2}} = \frac{\ln(n)}{n^{3/2}}.$$

另一方面, 当 n 很大时, 上述关系式两边都趋于 0. (要知道, 对数增长缓慢!) 所以可以运用关系当 $x \to 0$ 时 $\tan x \sim x$, 将 x 换成长串 $(n^2+4n-3)\ln(n)/\sqrt{n^7+2n^4+3n}$, 可得

$$\text{当 } n \to \infty \text{ 时,} \quad \tan\left(\frac{(n^2+4n-3)\ln(n)}{\sqrt{n^7+2n^4+3n}}\right) \sim \frac{(n^2+4n-3)\ln(n)}{\sqrt{n^7+2n^4+3n}} \sim \frac{\ln(n)}{n^{3/2}}.$$

现在我们来关注

$$\sum_{n=2}^{\infty} \frac{\ln(n)}{n^{3/2}}.$$

这里我们需要运用对数增长缓慢的事实, 即 $\ln(n)$ 较之 $n^{3/2}$ 不重要 (详见 21.3.4 节). 特别地, 我们需要分母中的幂次 3/2, 不希望其为 1 或更小. 故我们采用 $\ln(n) < Cn^{1/4}$(这里幂次仅需小于 1/2) 可得

$$\frac{\ln(n)}{n^{3/2}} \leqslant \frac{Cn^{1/4}}{n^{3/2}} = \frac{C}{n^{5/4}}.$$

因此, 把所有项加起来, 由 p 判别法可得

$$\sum_{n=2}^{\infty} \frac{\ln(n)}{n^{3/2}} \leqslant C \sum_{n=2}^{\infty} \frac{1}{n^{5/4}} < \infty.$$

所以由比较判别法可知

$$\sum_{n=2}^{\infty} \frac{\ln(n)}{n^{3/2}}$$

收敛. 现在回到前面的渐近关系, 由极限比较判别法可推得

$$\sum_{n=2}^{\infty} \tan\left(\frac{(n^2+4n-3)\ln(n)}{\sqrt{n^7+2n^4+3n}}\right)$$

也收敛. 太好了, 我们就要成功了. 那因子 $\cos^2(n)$ 呢? 这个因子不起什么作用, 因为它一直在振荡. 我们知道该因子小于等于 1, 且为正 (因为是平方). 所以我们只需看看由 $\cos^2(n) \leqslant 1$ 能得到什么. 事实上

$$\sum_{n=2}^{\infty} \cos^2(n) \tan\left(\frac{(n^2+4n-3)\ln(n)}{\sqrt{n^7+2n^4+3n}}\right) \leqslant \sum_{n=2}^{\infty} \tan\left(\frac{(n^2+4n-3)\ln(n)}{\sqrt{n^7+2n^4+3n}}\right) < \infty,$$

如我们刚刚得证的, 右边的级数收敛, 所以根据比较判别法可知原级数收敛. 在这个问题中, 我们用了两次比较判别法, 一次极限比较判别法和两次 p 判别法. 一堆令人迷惑的判别法. 不过如果你能够独立完成这类问题, 就能够解决任何一个涉及这三个判别法的问题了.

23.7 应对含负项的级数

对于某些项为负的级数, 这里有一些解决方法.

(1) **若所有项都为负, 则可通过在所有项前面加负号来修改级数:** 修改后的级数 $\sum\limits_{n=1}^{\infty}(-a_n)$ 所有项均为正. 然后运用前面所学的正项级数判别法来判定级数的敛散性. 若修改后的级数发散, 则原级数也发散; 若修改后的级数收敛, 则原级数也收敛. 事实上, 若修改后的级数收敛于 L, 则原级数收敛于 $-L$, 因为修改后的级数与原级数只相差一个负号. 例如, 级数

$$\sum_{n=3}^{\infty} \ln\left(\frac{1}{n}\right)\frac{1}{\sqrt{n}}$$

收敛还是发散? 当 n 很大时, $1/n$ 趋于 0, 所以它的对数值是一个负数. (要知道, 若 $0 < x < 1$, 则 $\ln(x) < 0$.) 所以现在考虑修改后的级数

$$\sum_{n=3}^{\infty} -\ln\left(\frac{1}{n}\right)\frac{1}{\sqrt{n}}$$

就比较容易了, 该级数其实与

$$\sum_{n=3}^{\infty} \ln(n)\frac{1}{\sqrt{n}}$$

是一样的, 因为 $-\ln(1/n) = -(\ln(1) - \ln(n)) = \ln(n)$. 有灵感了吗? 若该级数只是

$$\sum_{n=3}^{\infty} \frac{1}{\sqrt{n}},$$

由 p 判别法可知是发散的. 通常, 对数不起什么作用, 但这并不总是对的, 想一下前面的积分判别法的例题. 不管怎样, 这个特殊的对数帮助级数发散, 因为当 $n \to \infty$, 它无界. 基本的逻辑是 n 从 3 往上取值, $\ln(n)$ 的最小值是 $\ln(3)$, 故我们有

$$\ln(n) \geqslant \ln(3)$$

对任意 $n \geqslant 3$ 均成立. 在我们的级数中, 由 p 判别法 $(p = 1/2)$ 可得

$$\sum_{n=3}^{\infty} \ln(n)\frac{1}{\sqrt{n}} \geqslant \sum_{n=3}^{\infty} \frac{\ln(3)}{\sqrt{n}} = \ln(3)\sum_{n=3}^{\infty} \frac{1}{\sqrt{n}} = \infty,$$

即修改后的级数发散到 ∞, 由此可知原级数发散到 $-\infty$.

(2) **若有些项为正, 有些项为负, 尝试用第 n 项判别法:** 即, 验证当 $n \to \infty$ 时通项趋于 0, 否则, 马上可知级数发散. 例如,



$$\sum_{n=1}^{\infty} (-1)^n n^2$$

发散, 因为项 $(-1)^n n^2$ 的极限不为 0. (实际上, 极限不存在, 因为数列在越来越大的正数与负数之间来回振荡.) 这里没必要运用其他的判别法.

(3) **若有些项为正, 有些项为负, 且当 $n \to \infty$ 时通项趋于 0, 尝试应用绝对收敛判别法:**

$$\boxed{\text{若 } \sum_{n=1}^{\infty} |a_n| \text{ 收敛, 则 } \sum_{n=1}^{\infty} a_n \text{ 也收敛.}}$$

在这种情况下, 我们说数列*绝对收敛*. 例如, 级数

$$\sum_{n=1}^{\infty} \frac{\sin(n)}{n^2}$$

绝对收敛, 因为我们已经在 23.6 节见到

$$\sum_{n=1}^{\infty} \frac{|\sin(n)|}{n^2}$$

收敛. 所以对给定的有些项为正、有些项为负的级数, 若其不是明显不能用第 n 项判别法, 则需看一下该级数是否绝对收敛. 若该级数绝对收敛, 则其收敛; 若不是绝对收敛, 不要放弃, 继续下一步.

(4) **若级数不是绝对收敛, 尝试交错级数判别法:** 如 22.5.4 节所述,

$$\boxed{\text{若当 } n \to \infty \text{ 时交错级数的通项的绝对值单调递减趋于 0, 则级数收敛.}}$$

所以若想对级数 $\sum\limits_{n=1}^{\infty} a_n$ 应用该判别法, 有三件事情需要验证:

- a_n 的值在正负之间交替 (即各项的符号顺序为 $+, -, +, -, \cdots$, 或 $-, +, -, +, \cdots$);
- 随着 n 的增大, 量 $|a_n|$ 趋于 0, 即
$$\lim_{n\to\infty} |a_n| = 0;$$
- 绝对值 $|a_n|$ 关于 n 递减 (即通项的绝对值变得越来越小).

如果上面三个性质都满足, 则级数收敛. 注意: **无论何时都要首先尝试运用绝对收敛判别法. 若级数绝对收敛, 就不必用交错级数判别法!** 同样, 注意第二个性质是第 n 项判别法的另一种形式, 因为 $\lim\limits_{n\to\infty} |a_n| = 0$, 当且仅当 $\lim\limits_{n\to\infty} a_n = 0$. 所以, 即使你忘了先用第 n 项判别法, 但作为交错级数判别法的一种形式, 你还是会用到第 n 项判别法的.

这是一个经典的例子:
$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}.$$
根据 p 判别法, 其绝对值形式 $\sum\limits_{n=1}^{\infty} 1/n$ 发散. 所以原级数不是绝对收敛的. 现在直接应用交错级数判别法. 我们需要验证这三个性质. 首先, 级数交错吗? 是的. 一个级数如果含有形如 $(-1)^n$ 或 $(-1)^{n+1}$ 乘以一个正数的项, 则一定是交错的. 在这个例子中, 第 n 项是 $(-1)^n$ 与正数 $1/n$ 相乘. 那第二个性质呢? 我们需要证明
$$\lim_{n\to\infty} \left|\frac{(-1)^n}{n}\right| = 0.$$
此式显然成立, 因为 $|(-1)^n/n| = 1/n$. 对第三个性质, 我们需要证明 $\{|(-1)^n/n|\}$ 是一个递减数列. 可以很直接得出来, 还是因为 $|(-1)^n/n| = 1/n$, 且我们知道 $1/n$ 关于 n 递减. 所以可以应用交错级数判别法, 并得到原级数
$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$
收敛. 由于我们已经知道它不绝对收敛, 故其条件收敛.

另一方面, 考虑级数
$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}.$$
它的绝对值形式为 $\sum\limits_{n=1}^{\infty} 1/n^2$, 根据 p 判别法知其收敛. 所以上述级数绝对收敛, 因此没必要浪费时间在交错级数判别法上.

我们来看另一些例题. 首先来看

$$\sum_{n=1}^{\infty}(-1)^n \sin\left(\frac{1}{n}\right).$$

这与 23.6 节讲到的一个例题很像, 只是现在这个级数含有因子 $(-1)^n$. 对该级数要做的第一件事是验证其是否绝对收敛. 绝对值形式为

$$\sum_{n=1}^{\infty}\left|(-1)^n \sin\left(\frac{1}{n}\right)\right| = \sum_{n=1}^{\infty}\sin\left(\frac{1}{n}\right).$$

在原来的例题中, 我们已知当 $n \geqslant 1$, $\sin(1/n)$ 非负, 这就可以去掉绝对值号. 我们同样知道, 上式右侧级数发散, 所以原级数不是绝对收敛的. 另一方面, 该级数的各项显然交错, 且当 $n \to \infty$ 时通项趋于 0, 因为 $\sin(1/n)$ 是这样的. 现在考虑 $|a_n|$, 其实就是 $\sin(1/n)$. 它关于 n 递减吗? 对 $\sin(1/x)$ 关于 x 求导, 可得 $-\cos(1/x)/x^2$, 当 $x \geqslant 1$ 时它为负. 或者可以说, $1/n$ 关于 n 递减, 且在 0 附近 $\sin x$ 关于 x 递增, 所以 $\sin(1/n)$ 关于 n 递减. 不管用哪种方法, 我们已经证得了三个性质, 故由交错级数判别法可知级数收敛. 由于该级数不绝对收敛, 所以它条件收敛.

最后一个例题. 考虑级数

$$\sum_{n=1}^{\infty}(-1)^n\left(1+\frac{1}{n}\right)^n.$$

该级数显然交错, 但第 n 项的极限是多少? 若期望该级数收敛, 我们需要极限值为 0. 这里似乎有点问题, 根据 22.1.2 节末方框中的极限, 将 k 用 1 代换, 我们有

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \mathrm{e}^1 = \mathrm{e},$$

所以这个数列的交错形式在数 e 和 $-\mathrm{e}$ 之间振荡. 这意味着

$$\lim_{n\to\infty}(-1)^n\left(1+\frac{1}{n}\right)^n \text{ 不存在.}$$

由于极限不为 0, (甚至不存在!) 由第 n 项判别法知原级数

$$\sum_{n=1}^{\infty}(-1)^n\left(1+\frac{1}{n}\right)^n$$

肯定发散, 这里不要掉进交错级数判别法的陷阱, 否则会得出级数收敛的结论.

可见, 级数的讨论并不简单. 另外, 我们在下一章讨论幂级数和泰勒级数时仍会用到这些技术, 所以你非常有必要理解这一章的内容, 否则的话就难以应对后面接踵而来的问题. 当然, 大量做题会很有帮助.

第 24 章　泰勒多项式、泰勒级数和幂级数导论

现在我们讨论关于幂级数、泰勒多项式和泰勒级数的重要话题. 本章将从总体上探讨这些话题. 后面两章将讨论以本章为背景的解题方法. 下面是本章要讨论的内容：

- 近似值、泰勒多项式和泰勒近似定理;
- 近似值的精确度和完整的泰勒定理;
- 幂级数定义;
- 泰勒级数和麦克劳林级数定义;
- 泰勒级数的收敛性问题.

24.1　近似值和泰勒多项式

这里有个不错的结果：对任意实数 x, 我们有

$$\mathrm{e}^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{6}.$$

另外, x 越趋近于 0, 近似程度就越好.

现在我们来讨论这个结果. $x = 0$ 时, 两边其实都等于 1, 所以这个近似值很理想! 那 x 不为 0 时呢? 我们试一下 $x = -1/10$, 由上述等式可得

$$\mathrm{e}^{-1/10} \approx 1 - \frac{1}{10} + \frac{1/100}{2} - \frac{1/1000}{6},$$

化简可得

$$\mathrm{e}^{-1/10} \approx \frac{5429}{6000}.$$

根据计算器所得结果, $\mathrm{e}^{-1/10}$ 等于 0.904 837 418 0 (精确到 10 位小数), 而 5429/6000 等于 0.904 833 333 3 (也精确到 10 位小数). 这些数很接近. 事实上, 它们的差仅为 0.000 004 084 7.

那到底我是怎么想出多项式 $1 + x + x^2/2 + x^3/6$ 的呢? 很显然, 它不只是一个旧多项式, 特别的是它与 e^x 有关系. 与其只关注 e^x, 倒不如考虑其他更一般的函数. 同样, 该多项式的次数 3 也没有什么特别的, 我们可用任何次数. 我们就从次数 1 开始吧, 看看会发生什么.

24.1.1　重访线性化

我们说, 有些光滑函数 f 可以被求任意阶导而不会出现任何问题. 这里有个

13.2 节问过的问题：在点 $(a, f(a))$ 附近, 与曲线 $y = f(x)$ 最近似的直线方程是什么？答案是所求直线为曲线上点 $(a, f(a))$ 处的切线, 它的方程为

$$y = f(a) + f'(a)(x - a).$$

这就是 f 在 $x = a$ 的线性化. 右边是次数为 1 的多项式. 图 24-1 给出了曲线 $y = f(x)$ 在 $x = a$ 的切线, 看起来不像是整个曲线的近似.

不过, 它确实为曲线在 $(a, f(a))$ 附近的近似. 事实上, 我们把 $(a, f(a))$ 附近放大, 如图 24-2 所示. 可以看到, 切线与曲线 $y = f(x)$ 并没有很大的差别. 图像放得越大, 它们的差别就越小.

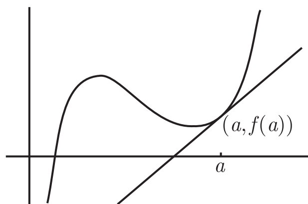


图 24-1

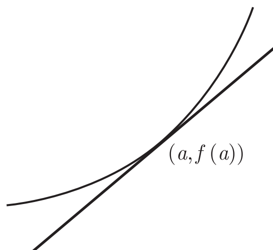


图 24-2

24.1.2 二次近似

为什么只讨论直线？我们再来探讨这个与前一节开始相同的问题, 但这次讨论抛物线. 问题：在点 $(a, f(a))$ 附近, 与曲线 $y = f(x)$ 最近似的二次曲线方程是什么？采用相同的函数, 图 24-3 是我们猜出的二次曲线可能的样子.

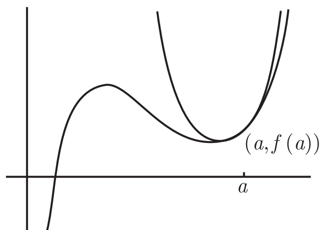


图 24-3

事实上, 在 x 接近于 a 时 (即, 曲线上点 $(a, f(a))$ 附近), 最近似于曲线 $y = f(x)$ 的二次曲线方程为

$$y = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2.$$

它其实是一个关于 x 的二次函数, 因为若展开 $(x - a)^2$, 则 x 的最高次项为 x^2. 这里仍保留了相同的形式, 并称之为 “关于 $(x - a)$ 的二次函数”. 我们称该二次函数为 P_2, 即

$$P_2(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2.$$

现在, 我们搜集一些关于 P_2 的好结论.

(1) 将 $x = a$ 代入方程 $P_2(x)$, 可以很容易地得到 $P_2(a) = f(a)$. 所以当 $x = a$ 时, P_2 与 f 的值相等. 事实上, 因为函数的零阶导为该函数本身, 所以当 $x = a$ 时, P_2 与 f 的零阶导相等.

(2) 对 P_2 求导可得 $P_2'(x) = f'(a) + f''(a)(x-a)$. 同样, 若代入 $x = a$, 可知 $P_2'(a) = f'(a)$. 当 $x = a$ 时, 一阶导 P_2' 与 f' 也相等.

(3) 再求一次导可得 $P_2''(x) = f''(a)$. 当 $x = a$, 有 $P_2''(a) = f''(a)$. 所以当 $x = a$, 二阶导数值也相等.

(4) 另一方面, 由于 $f''(a)$ 为常数, 所以对所有 x 有 $P_2'''(x) = 0$. 对所有更高阶导数均有相同结论. (毕竟, P_2 是二次的, **任何**二次函数的三阶或更高阶导数必处处为 0!)

所以 P_2 与 f 在 $x = a$ 有相同的零阶导、一阶导和二阶导, 但 P_2 的三阶或更高阶导恒为 0. 可以说, P_2 提取了 f 在 $x = a$ 处二阶导及以下的所有信息.

另一个关于 P_2 的好结论是：若忽视 $P_2(x)$ 方程右边的最后一项, 就得到 $f(a)+f'(a)(x-a)$. 这恰恰是上一节的线性化, 所以可以认为最后的项 $\frac{1}{2}f''(a)(x-a)^2$ 为所谓的**二阶修正项**. 这意味着我们应该能够找到比切线更好的近似. 二阶修正项有助于更接近于曲线, 至少当 x 在 a 附近时是这样的. (当 $f''(a) = 0$ 时是例外, 在这种情况下 P_2 仅为线性化, 并不能使近似更好.)

24.1.3　高阶近似

我们继续相同形式的讨论, 只不过这里用任意次 N 代替 1 或 2. 问题：对 a 附近的 x, 哪个次数为 N 或更低的多项式最近似于 $f(x)$? 答案由下面的定理给出.

> **泰勒近似定理：** 若 f 在 $x = a$ 光滑, 则在所有次数为 N 或更低的多项式中, 当 x 在 a 附近时, 最近似于 $f(x)$ 的是
>
> $$P_N(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(N)}(a)}{N!}(x-a)^N.$$

用求和号表示该公式为：

$$P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

在这个公式中, 要知道 $0! = 1$, $f^{(0)}(a)$ 与 $f(a)$ 意思一样 (零阶导数), $f^{(1)}(a)$ 与 $f'(a)$ 意思一样 (一阶导数).

我们称多项式 P_N 为 $f(x)$ 在 $x = a$ 处的 N **阶泰勒多项式**. 注意 P_N 的次数可能小于 N. 例如, 若 $f^{(N)}(a) = 0$, 则上述和式的最后一项为 0, P_N 的次数至多为 $N-1$. 因此, 我们称之为 N 阶泰勒多项式而不是 N 次泰勒多项式. (多项式 $P_N(x)$ 有时被写成 $P_N(x; a)$, 以强调每次选择不同的 N 和 a 得到不同的多项式. 我将采用形式 $P_N(x)$, 因为每次讨论我们只选择一个 a.)

再次强调, P_N 的重要性质是对所有 $n=0,1,\cdots,N$,

$$P_N^{(n)}(a)=f^{(n)}(a),$$

即当 $x=a$ 时, P_N 的所有 N 阶导及以下的导数值都与 f 对应值相等, 但是 P_N 的所有更高阶导数必须处处为 0. 函数 P_N 提取了 f 在 $x=a$ 处直到 N 阶导数的所有信息.

当然, 当 $N=1$ 时, 我们得到 $P_1(x)=f(a)+f'(a)(x-a)$, 为 f 在 $x=a$ 处的线性化. 当 $N=2$ 时, 我们就得到上一节的公式 $P_2(x)$. 下面看一下该方法对 $a=0$ 的 $f(x)=\mathrm{e}^x$ 的应用. 由上面的公式, 令 $N=3$ 且 $a=0$, 我们有

$$P_3(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\frac{f^{(3)}(0)}{3!}x^3.$$

幸运的是, e^x 关于 x 的所有导数均为 e^x, 所以可知 $f(0)$, $f'(0)$, $f''(0)$ 和 $f^{(3)}(0)$ 都是 e^0, 等于 1. 由于 $2!=2$, $3!=6$, 因而上述公式变为

$$P_3(x)=1+x+\frac{1}{2}x^2+\frac{1}{6}x^3.$$

这恰恰是 24.1 节开头提到的三次多项式! 在所有的次数为 3 或更低次的多项式中, 这个多项式是与 x 在 0 附近的 e^x 最近似的. 为什么是 0 呢? 因为那是我们所选择的 a 值. 若选择不同的 a 值, 我们将得到 x 在 a 附近对 e^x 有很好近似的另一个不同的多项式. 去掉三次项 $x^3/6$ 后可以看到, $P_2(x)=1+x+x^2/2$, 然后再去掉二次项 $x^2/2$, 得到线性化 $P_1(x)=1+x$. 从另一个角度来看, $P_2(x)$ 通过加上二阶修正项 $x^2/2$ 而改进了 $P_1(x)$, 而 $P_3(x)$ 通过加上三阶修正项 $x^3/6$ 而改进了 $P_2(x)$. 每次使 N 加 1 都会使近似通过加上另一个修正项而变得更好.

其实泰勒近似定理依赖于泰勒定理, 泰勒定理将在下一节讨论. 近似定理也有一些模糊不清的说法:"最好的近似" 究竟是什么意思? 我们将在下一节进一步探讨, 真正的答案连同定理证明在附录 A.7 节.

24.1.4 泰勒定理

在 24.1 节开始, 我们看到

$$\mathrm{e}^x\approx 1+x+\frac{x^2}{2}+\frac{x^3}{6}.$$

特别地, 注意当 $x=-1/10$, 上面的近似变为了

$$\mathrm{e}^{-1/10}\approx 1-\frac{1}{10}+\frac{1/100}{2}-\frac{1/1000}{6}=\frac{5429}{6000}.$$

这个近似有多好? 衡量的一个方法是, 考虑真正的量 $\mathrm{e}^{-1/10}$ 与近似值 $5429/6000$ 的差. 我们称这个差量为近似的*误差*, 因为它指出了用近似值代替真实值的错误有多大. 该例子的误差是:

$$误差 = 真实值 - 近似值 = e^{-1/10} - \frac{5429}{6000}.$$

若误差很小, 则近似程度较好. 在 24.1 节, 我们看到近似到 10 位小数时差值为 0.000 004 084 7, 但我们需要用计算器, 且这不是我们做近似的全部目的. 要知道, 计算器给出的数也是近似值! 此外, 你认为计算器是怎么工作的? 可能它是用泰勒多项式求出 $e^{-1/10}$ 的近似的.

我们真正喜欢的是误差的另一个公式, 泰勒定理由此而来. 与其只讨论特定的例子 e^x, 倒不如讨论更一般的问题. 我们正在讨论光滑函数 f 和它的关于 $x = a$ 的 N 阶泰勒多项式, 如前一节所见, 该多项式为

$$P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

我们想用 $P_N(x)$ 的值来获取 $f(x)$ 的近似值, 所以考虑误差项, 即真实值和近似值之差:

$$R_N(x) = f(x) - P_N(x).$$

实际上, $R_N(x)$ 被称为N *阶误差项*, 也称为 N *阶余项*, 因为它就是从 $f(x)$ 取走 $P_N(x)$ 所余下的部分. 如前所述, 泰勒定理给出了 $R_N(x)$ 的另一个公式:

泰勒定理: 关于 $x = a$ 的 N 阶余项

$$R_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!}(x-a)^{N+1},$$

其中 c 是介于 x 与 a 之间的一个数.

注意数 c 依赖于 x 和 N, 一般不能确定! 由于 $f(x) = P_N(x) + R_N(x)$, 则我们可以写为

$$f(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(N+1)}(c)}{(N+1)!}(x-a)^{N+1}.$$

这个式子看起来很不舒服. 而且这个 c 究竟是什么呢? 其实, 我们以前见过类似的情况. 回顾 11.3 节对中值定理 (MVT) 的讨论. 由 MVT, 若 f 在区间 $[a, b]$ 上足够光滑, 则在 $[a, b]$ 上存在一个数 c(其值一般不能确定), 使得

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

若将 b 换为 x, 并解出 $f(x)$, 可得

$$f(x) = f(a) + f'(c)(x-a),$$

其中 c 介于 a 和 x 之间. 现在回到泰勒定理那个最后的等式并令 $N = 0$. $P_0(x)$ 是什么? 就是 $f(a)$. 那 $R_0(x)$ 呢? 根据泰勒定理,

$$R_0(x) = \frac{f^{(1)}(c)}{1!}(x-a)^1 = f'(c)(x-a),$$

其中 c 介于 a 和 x 之间. 则泰勒定理 $(N=0)$ 有

$$f(x) = P_0(x) + R_0(x) = f(a) + f'(c)(x-a),$$

这就是 MVT 的内容! 所以, 泰勒定理基本上是中值定理的扩展. 另外, 这里说 c 介于 a 和 x 之间而不是 $a \leqslant c \leqslant x$, 是因为 x 也可能比 a 小, 那样的话, 我们将会有 $x \leqslant c \leqslant a$.

现在令 $N=1$ 而不是 $N=0$. 上面方框中的主要公式变为

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(c)}{2!}(x-a)^2 = L(x) + R_1(x);$$

这里 $L(x) = f(a)+f'(a)(x-a)$ 是 f 关于 $x=a$ 的线性化, 且 $R_1(x) = \frac{1}{2}f''(c)(x-a)^2$ 为一阶误差项. 这与我们在 13.2.4 节给出的误差项 $r(x)$ 一致.

回到 e^x 的近似, 当我们写

$$\mathrm{e}^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

时, 就是在说 $\mathrm{e}^x \approx P_3(x)$, 其中 P_3 是 $f(x) = \mathrm{e}^x$ 关于 $x=0$ 的三阶泰勒多项式. 由泰勒定理, R_3 为

$$R_3(x) = \frac{f^{(4)}(c)}{4!}x^4,$$

其中 c 介于 0 和 x 之间. (我只是将 $N=3$ 和 $a=0$ 代入上面方框中的 $R_N(x)$ 公式.) 由于 e^x 的任意阶导数 (关于 x 的) 都为 e^x, 我们可以知道 $f^{(4)}(c) = \mathrm{e}^c$, $4! = 24$, 所以有

$$R_3(x) = \frac{\mathrm{e}^c}{24}x^4.$$

换句话说,

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{\mathrm{e}^c}{24}x^4.$$

我们已经把近似变成一个方程, 但不知道 c 的值! 不过我们还是从这里得到了一些有用的东西, 因为我们知道 c 介于 0 和 x 之间. 例如, 若再次令 $x = -1/10$, 可得

$$\mathrm{e}^{-1/10} = 1 - \frac{1}{10} + \frac{1/100}{2} - \frac{1/1000}{6} + \frac{\mathrm{e}^c}{24}(1/10\,000),$$

上式可化简为

$$\mathrm{e}^{-1/10} = \frac{5429}{6000} + \frac{\mathrm{e}^c}{240\,000}.$$

这次, 我们知道 c 介于 0 和 $x = -1/10$ 之间, 所以其实有 $-1/10 < c < 0$. 因为 e^c 关于 c 递增, 显然 c 足够大, e^c 就会最大. 这就意味着 $c=0$ 时 e^c 有最大值, 这样 e^c 就不会比 $\mathrm{e}^0 = 1$ 大. 所以误差项至多为 $1/240\,000$. 换句话说, 当写 $\mathrm{e}^{-1/10} \approx 5429/6000$

时, 我们知道近似的精确度要好于 1/240 000, 大约为 0.000 004 166 7. (将该值与 24.1 节的差的实际值比较一下.)

我们将在 25.3 节看一些运用泰勒定理的例子. 现在来讲讲幂级数和泰勒级数.

24.2　幂级数和泰勒级数

这是另一个结论：

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

对所有实数 x 均成立. 你可能会注意到, 它与 24.1 节开头的近似类似, 但有两点明显不同. 首先, 我们不再讨论近似; 其次, 右边是一个无穷级数. 当面对无穷级数时, 要小心了.

我们来看一下, 能否理解上述等式的意义. 假定从

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

开始, 它看起来像是一个多项式, 但实际上不是, 因为没有最高次项. 它只是一直继续下去. 其实, 它是一个幂级数. 若将 x 换成任意一个特定的值, 就得到一个常规的级数. 例如, 若 $x = -1/10$, 得到级数

$$1 - \frac{1}{10} + \frac{1/100}{2!} - \frac{1/1000}{3!} + \frac{1/10\,000}{4!} - \frac{1/100\,000}{5!} + \cdots,$$

也可以另写为

$$1 - \frac{1}{10} + \frac{1}{100 \times 2!} - \frac{1}{1000 \times 3!} + \frac{1}{10\,000 \times 4!} - \frac{1}{100\,000 \times 5!} + \cdots.$$

该级数可能收敛, 也可能发散. 那到底是收敛还是发散呢？答案是收敛. 还有, 我们甚至知道它收敛于 $e^{-1/10}$. 这就使我们知道了, 对任意实数 x,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

都成立. 意思是若将 x 的任何特定值代入右边, 就得到一个收敛于 e^x 的级数. 我们将在 24.2.3 节证明这个结论的正确性. 从下面的例子可以看出, 代入 x 的一些不同值会得到什么：

$$x = 2 : 1 + 2 + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!} + \frac{2^5}{5!} + \cdots \qquad \text{收敛于 } e^2$$

$$x = -5 : 1 - 5 + \frac{5^2}{2!} - \frac{5^3}{3!} + \frac{5^4}{4!} - \frac{5^5}{5!} + \cdots \qquad \text{收敛于 } e^{-5}$$

$$x = 0 : 1 + 0 + 0 + 0 + 0 + \cdots \qquad \text{收敛于 } 1$$

我还可以给出更多例子, 实际上有无穷多个. 这个幂级数给出了无穷多个常规级数的信息, 一个 x 值对应一个级数. 显然, 上面最后的级数收敛于 1. 令 $x = 0$, 会发生

很特别的事情：它使得除了常数项外的其他项都消失了. 我们很快就会讨论这点, 先来看一般的幂级数.

24.2.1 一般幂级数

关于 $x=0$ 的幂级数是形为

$$a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+\cdots$$

的式子, 其中数 a_n 是确定的常数. 尽管幂级数不是一个多项式, 我们仍可定义 a_n 为幂级数中 x^n 的系数. 上述级数也可以用求和号写为

$$\sum_{n=0}^{\infty}a_nx^n.$$

在前一节的例子中, 相应级数是

$$1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\cdots,$$

用求和号可写成

$$\sum_{n=0}^{\infty}\frac{1}{n!}x^n.$$

所以这是一个系数定义为 $a_n=1/n!$ 的幂级数, 其中 n 为任意非负整数. 注意, x 是唯一的变量, n 只不过是一个虚拟变量, 一旦将和式展开它就消失了. 比上述幂级数更简单的一个级数的展开式及求和号表示形式为

$$1+x+x^2+x^3+x^4+\cdots=\sum_{n=0}^{\infty}x^n.$$

在这个级数中, 系数 a_n 都等于 1. 希望你能够看出这是首项为 1、公比为 x 的几何级数.

对给定的 x, 我们经常将方程写成

$$f(x)=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+\cdots$$

的形式. 意思是若将 x 取值范围内的一个值代入, 幂级数就变为收敛于值 $f(x)$ 的常规级数. 例如, 我们已经说过 (但未证明) 的

$$\mathrm{e}^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\cdots$$

对所有 x 均成立. 另一方面, 当我们讨论 22.2 节等比数列的求和方法时, 有

$$1+r+r^2+r^3+r^4+\cdots=\sum_{n=0}^{\infty}r^n=\frac{1}{1-r},\quad -1<r<1.$$

用 x 代换 r：

$$1+x+x^2+x^3+x^4+\cdots=\sum_{n=0}^{\infty}x^n=\frac{1}{1-x},\quad -1<x<1.$$

即, 当 $-1<x<1$,

$$\frac{1}{1-x}=1+x+x^2+x^3+x^4+\cdots.$$

若将 x 换成任意一个在该区间的数, 右边就得到一个常规级数, 收敛于左边的值. 另一方面, 若 $x>1$ 或 $x\leqslant 1$ 又会如何呢? 左边有意义, 但右边没有意义, 因为对 x 的这些值, 级数发散. (当 x 等于 1 时, 两边都无定义.)

当 $x=0$ 时, 幂级数

$$a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+\cdots$$

有一些很好的性质：除了开始的 a_0, 其他所有项都没了, 所以级数自动收敛. (当然, 收敛于 a_0!) 但这并没有告诉我们, 对其他的 x 值, 级数是否收敛. 例如, 几何级数只有当 $-1<x<1$ 时收敛, 而我们将在 26.1.2 节指出, 下面的级数只有当 $x=0$ 时收敛：

$$\sum_{n=0}^{\infty}n!x^n$$

不可否认, 0 是一个很受欢迎的数, 但它并不比其他的实数特殊. 我们可将这个特殊的性质转移到其他的数 a. 我们只需将 x 换为 $(x-a)$. 故下面是幂级数在 $x=a$ 的一般表达式：

$$\boxed{a_0+a_1(x-a)+a_2(x-a)^2+a_3(x-a)^3+a_4(x-a)^4+\cdots.}$$

用求和号表示为

$$\boxed{\sum_{n=0}^{\infty}a_n(x-a)^n.}$$

当 $x=a$ 时, 该级数当然收敛, 因为除了 a_0 外, 其他所有项都没有了. 数 a 称为幂级数的中心. 什么时候需要考虑中心不为 0 的幂级数呢? 一个可能的例子是, 你想求收敛于 $\ln(x)$ 的幂级数. 该量在 $x=0$ 没有定义, 所以想求收敛于 $\ln(x)$ 的在 $x=0$ 的幂级数是愚蠢的行为. 另一方面, 我们能找到一个以 1 为中心且收敛于 $\ln(x)$ 的幂级数, 至少对 x 的某些值是可以的. 实际上, 在 26.2.1 节, 我们将看到等式

$$\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}(x-1)^n=\ln(x)$$

对 $-1<(x-1)<1$ 成立, 即对 $0<x<2$ 成立. (甚至对 $x=2$ 也成立：

$$\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\cdots=\ln(2).$$

不过, 这个不是很容易证明!)

24.2.2 泰勒级数和麦克劳林级数

在上一节中, 我们看到在 $x = a$ 时的一般幂级数为 (求和号表示形式和展开式)

$$\sum_{n=0}^{\infty} a_n(x-a)^n = a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + a_4(x-a)^4 + \cdots.$$

在 $x = a$ 收敛, 也可能在 x 的其他值收敛. 在 26.1.2 节, 我们将讨论求使级数收敛的 x 值的方法. 我们可以每次代换一个 x 值, 看看每种情况下, x 收敛于何值, 并称收敛的值为 $f(x)$. 我们从幂级数开始, 需定义一个函数.

假设我们从一个光滑函数 f 开始. 用 f 的所有导数定义一个在 $x = a$ 的幂级数

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

将求和号展开后变为

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \frac{f^{(4)}(a)}{4!}(x-a)^4 + \cdots.$$

该幂级数的系数为 $a_n = f^{(n)}(a)/n!$. 该级数称为 f *关于* $x = a$ *的泰勒级数*. 所以, 从函数开始, 我们定义了幂级数.

仔细看一下上面泰勒级数的定义, 应该很面熟吧. 其实, 该公式与 24.1.3 节中泰勒多项式 $P_N(x)$ 的定义很像. 唯一的区别是, 和式没有终止于 $n = N$, 而是一直持续到 ∞. 换句话说, 泰勒多项式 $P_N(x)$ 是泰勒级数的 N 项部分和.

我们将在下一节讨论泰勒多项式和泰勒级数的联系. 首先, 我们有另一个定义: *麦克劳林级数*, 它是 f 关于 $x = 0$ 的泰勒级数的另一个名字. 所以,

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n,$$

展开式为

$$f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdots.$$

无论什么时候看到 "麦克劳林级数" 这几个字, 脑子里想着 "$a = 0$ 的泰勒级数" 就可以了.

24.2.3 泰勒级数的收敛性

好, 我们来回顾一下. 我们从一个函数 f 和数 a 开始, 构造了 f 关于 $x = a$ 的泰勒级数:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

这是一个中心为 a 的幂级数, 但不仅仅是旧幂级数：它包含了 f 在 $x=a$ 的所有导数值. 若能写为

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n$$

将会很酷, 因为那样我们就会知道泰勒级数对任何 x 都收敛, 且收敛于原函数值 $f(x)$. 问题是, 上面的等式并不总是成立. 级数可能会对 x 的某些值发散, 或者对所有 x 值都发散 (除了 $x=a$：如我们看到的, 幂级数在它的中心总收敛). 更糟的是, 级数可能收敛于不是 $f(x)$ 的某些值! 幸运的是, 我们在例子中将避开这种离奇的可能性①.

那么, 你是怎么知道泰勒级数是否且何时收敛于原来的函数呢? 跟 24.1.4 节一样, 从

$$f(x)=P_N(x)+R_N(x)$$

开始. 记住,

$$P_N(x)=\sum_{n=0}^{N}\frac{f^{(n)}(a)}{n!}(x-a)^n \quad 和 \quad R_N(x)=\frac{f^{(N+1)}(c)}{(N+1)!}(x-a)^{N+1}.$$

这将 $f(x)$ 表示为近似值 $P_N(x)$ 与误差或者余项 $R_N(x)$ 的和. 这里比较聪明的想法是：令 N 越来越大. 这样就有希望使近似值 $P_N(x)$ 越来越接近于实际值 $f(x)$; 也就是希望误差 $R_N(x)$ 越来越小.

我们尝试用等式来描述上面的论述. 假设对某些 x, 我们已知

$$\lim_{N\to\infty}R_N(x)=0.$$

对等式 $f(x)=P_N(x)+R_N(x)$ 取 $N\to\infty$ 的极限, 得到

$$\lim_{N\to\infty}f(x)=\lim_{N\to\infty}P_N(x)+\lim_{N\to\infty}R_N(x)=\lim_{N\to\infty}P_N(x).$$

由于 $f(x)$ 不依赖 N, 左边就是 $f(x)$, 所以我们知道

$$f(x)=\lim_{N\to\infty}P_N(x)=\lim_{N\to\infty}\sum_{n=0}^{N}\frac{f^{(n)}}{n!}(x-a)^n=\sum_{n=0}^{\infty}\frac{f^{(n)}}{n!}(x-a)^n.$$

所以 $f(x)$ 等于它的泰勒级数! 换句话说, **若想证明一个函数在某些数 $\boldsymbol{x}$ 处等于它的泰勒级数, 可尝试证明当 $\boldsymbol{N\to\infty}$ 时 $\boldsymbol{R_N(x)\to 0}$**.



我们对 $f(x)=\mathrm{e}^x$, $a=0$ 做这些讨论. 通过改动 24.1.4 节提到的一些结论, 你应该可知

① 我只提及一个属于这种泰勒级数的经典例子：若 $f(x)=\mathrm{e}^{-1/x^2}$, 当 $x\neq 0$, 同时我们定义 $f(0)=0$, 则 f 在 0 点的所有导数均为 0, 所以 f 在中心 0 点的泰勒级数为 0. 除了当 $x=0$ 外, 这个泰勒级数与 $f(x)$ 一点都不同.

$$P_N(x) = \sum_{n=0}^{N} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^N}{N!},$$

以及对于介于 x 和 0 之间的某些 c,

$$R_N(x) = \frac{\mathrm{e}^c}{(N+1)!} x^{N+1}.$$

现在我们需要求 $R_N(x)$ 当 $N \to \infty$ 的极限并说明该极限为 0:

$$\lim_{N\to\infty} R_N(x) = \lim_{N\to\infty} \mathrm{e}^c \frac{x^{N+1}}{(N+1)!}.$$

在 24.3 节, 我将证明

$$\lim_{N\to\infty} \frac{x^{N+1}}{(N+1)!} = 0$$

对任意 x 成立. 我们要对因子 e^c 多加小心, 因为它依赖于 N. 问题是, e^c 会有多大? 要知道 c 介于 x 和 0 之间. 若 x 为负, e^c 的最大值可能出现在 $c = 0$, 意味着 $\mathrm{e}^c \leqslant 1$; 若 x 为正, e^c 的最大值可能出现 $c = x$, 意味着 $\mathrm{e}^c \leqslant \mathrm{e}^x$. 不管是哪种情况, 因为 x 是固定的 (即, 看作常数), 我们可以有 $0 \leqslant \mathrm{e}^c \leqslant C$, 其中 C 是另一个常数. 无论 N 为何值都成立, 即便 c 随着 N 的改变而在 x 和 0 之间变动. 不管怎样, 希望你相信这些, 由此你就会相信

$$0 \leqslant \mathrm{e}^c \frac{|x|^{N+1}}{(N+1)!} \leqslant C \frac{|x|^{N+1}}{(N+1)!}.$$

现在左边和右边都随着 N 趋于 ∞ 而趋于 0, 所以由三明治定理知, 中间的量也趋于 0. 我们已然证明了

$$\lim_{N\to\infty} R_N(x) = 0$$

对任意实数 x 成立. 这就意味着我们最终证明了

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

对所有实数 x 成立.

我们通过求 $f(x) = \cos(x)$ 的麦克劳林级数并证明它对所有 x 都收敛于 $f(x)$, 来讨论一下详细过程. 首先需要对 f 连续求导, 然后将 0 代入每个导数看一下会发生什么. 当对 $\cos(x)$ 关于 x 连续求导时, 得到 $-\sin(x)$, 然后是 $-\cos(x)$, 之后重复出现 $\sin(x), \cos(x), -\sin(x), -\cos(x), \cdots$, 且显然会循环下去. 当将 $x = 0$ 代入时, $\sin(x)$ 项没了, $\pm\cos(x)$ 项变为 ± 1, 所以数列 $f^{(n)}(0)$ 为



$$1, 0, -1, 0, 1, 0, -1, 0, 1, 0, -1, 0, \cdots.$$

若将这些数代入麦克劳林公式, 得到

$$f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \frac{f^{(5)}(0)}{5!}x^5 + \frac{f^{(6)}(0)}{6!}x^6 + \cdots,$$

所有奇次项都没有了, 即

$$1-\frac{1}{2!}x^2+\frac{1}{4!}x^4-\frac{1}{6!}x^6+\cdots,$$

写为更紧凑的形式为

$$1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots.$$

这就是 $\cos(x)$ 的麦克劳林级数, 或者称为 $\cos(x)$ 关于 $x=0$ 的泰勒级数. 为了得到相应的泰勒多项式, 所需做的就是削减级数右边. 例如,

$$P_4(x)=1-\frac{1}{2!}x^2+\frac{1}{4!}x^4.$$

顺便说一下, $P_5(x)$ 的公式与 $P_4(x)$ 公式是一样的, 因为上面的麦克劳林级数没有 5 次项. 这就说明了我们为什么要用 "阶" 这个词: P_5 的阶为 5, 但次数为 4.

剩下需要证明的是对所有的实数 x, $\cos(x)$ 都等于它的麦克劳林级数:

$$\cos(x)=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots.$$

为此, 我们要证明

$$\lim_{N\to\infty}R_N(x)=0.$$

我们知道

$$R_N(x)=\frac{f^{(N+1)}(c)}{(N+1)!}x^{N+1},$$

其中 c 介于 x 和 0 之间. 取绝对值:

$$|R_N(x)|=\frac{|f^{(N+1)}(c)|}{(N+1)!}|x|^{N+1}.$$

f 的所有导数或者为 $\pm\cos(x)$, 或者为 $\pm\sin(x)$, 所以 $\left|f^{(N+1)}(c)\right|$ 或者为 $|\cos(c)|$, 或者为 $|\sin(c)|$. 在任一种情况下, 这个量都小于等于 1, 所以我们有

$$0\leqslant|R_N(x)|\leqslant\frac{1}{(N+1)!}|x|^{N+1}.$$

在下一节, 我们还将证明

$$\lim_{N\to\infty}\frac{x^{N+1}}{(N+1)!}=0.$$

现在用三明治定理来证明

$$\lim_{N\to\infty}|R_N(x)|=0,$$

这同样意味着

$$\lim_{N\to\infty}R_N(x)=0.$$

我们已经证明了

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

对所有实数 x 都成立. 我们把上述级数用求和号来表示吧. (什么, 这不是你对解决难题的庆祝方式吗？) 不管怎样, 你是怎么只得到 x 的偶次幂的？答案是用 $2n$ 替换 n(见 15.1 节对这类问题的讨论). 由于分母上的阶乘与次数一致, 我们猜测, 该麦克劳林级数可写为

$$\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}.$$

问题是这个级数不是交错级数, 所以需要插入一个因子 $(-1)^n$：

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.$$

若将其展开, 你会发现这个改变是对的, 即

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

对所有实数 x 成立.

24.3 一个有用的极限

这一节跟幂级数没有关系, 只是关于前面几节中用过两次的极限的证明：

$$\lim_{N\to\infty} \frac{x^{N+1}}{(N+1)!} = 0$$

对所有实数 x 成立. 令 $n = N + 1$ (就像积分中的换元), 则与证明

$$\lim_{n\to\infty} \frac{x^n}{n!} = 0$$

对所有实数 x 成立一样. 有一些方法可证明后一个结论, 不过这有一个不显眼的方法. 我先来解释一下要用到的逻辑, 然后再用该方法. 我将证明级数

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

收敛, 不必考虑 x 是什么. (是的, 我们 “知道” 它其实收敛于 e^x, 但这是等到我们证明极限为 0 之后才知道的!) 不管怎样, 级数收敛于什么没关系, 仅仅知道级数收敛就足够了. 为什么？因为那样的话, 第 n 项 $x^n/n!$ 一定随着 n 趋于 ∞ 而趋于 0, 否则第 n 项判别法就不对了. 即, 若通项随着 n 趋于 ∞ 不趋于 0, 则级数将发散. 因此我们用比式判别法来证明级数对所有 x 收敛. 固定 x, $a_n = x^n/n!$, 看一下比值的极限：

$$L = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{x^{n+1}/(n+1)!}{x^n/n!}\right| = \lim_{n\to\infty}\left|\frac{x^{n+1}}{x^n}\frac{n!}{(n+1)!}\right|.$$

我们知道 $n!/(n+1)!$ 可化简为 $1/(n+1)$, 所以最后的极限为

$$\lim_{n\to\infty}|x|\frac{1}{n+1} = 0,$$

由于 $|x|$ 固定且 $1/(n+1)$ 趋于 0, 极限为 0, 小于 1, 所以级数收敛, 且我们也顺便证明了极限的正确性. 固定 x, 然后对该特定的 x 运用比式判别法, 来判别级数收敛的方法将在 26.1.2 节多次用到.

第 25 章　求解估算问题

在上一章中, 我们学习了如何应用泰勒多项式来估算 (或近似) 特定的量. 我们也知道了, 余项可以用来判定近似程度. 本章, 我们将详述相应的方法并讨论一些相关例题. 本章的计划是：

- 泰勒多项式和泰勒级数的重要结论回顾;
- 如何求泰勒多项式和泰勒级数;
- 估算问题;
- 分析误差的一个不同的方法.

25.1　泰勒多项式与泰勒级数总结

下面是关于泰勒多项式和泰勒级数的一些重要结论, 已在前一章中讨论过.

(1) 在所有次数为 N 或更低的多项式中, 与定义在 a 附近的光滑函数 f 最近似的多项式被称为关于 $x=a$ 的 N 阶泰勒多项式, 即

$$\boxed{\begin{aligned}P_N(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 \\ + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(N)}(a)}{N!}(x-a)^N.\end{aligned}}$$

用求和号表示, 可写为

$$\boxed{P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n.}$$

(2) 多项式 P_N 与 f 在 $x=a$ 点直到 N 阶的导数相同. 即

$$P_N(a) = f(a), \quad P_N'(a) = f'(a), \quad P_N''(a) = f''(a), \quad P_N^{(3)}(a) = f^{(3)}(a),$$

且直到 $P_N^{(N)}(a) = f^{(N)}(a)$. 一般来说, 对 a 之外的其他任何值, 或大于 N 的任何阶导数, 上述等式都不成立. (实际上, P_N 的大于 N 阶的所有导数都等于 0, 因为 P_N 是次数为 N 的多项式.)

(3) N 阶余项$R_N(x)$, 或称为 N 阶误差项, 是 $f(x) - P_N(x)$. 对任意 N 有

$$\boxed{f(x) = P_N(x) + R_N(x).}$$

余项表达式为

$$\boxed{R_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!}(x-a)^{N+1},}$$

其中 c 一般求不出来, 它介于 x 与 a 之间.

(4) 所以, $f(x)$ 的完整表达式为

$$\boxed{f(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(N+1)}(c)}{(N+1)!}(x-a)^{N+1}.}$$

(5) 无穷级数

$$\boxed{\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n}$$

被称为 $f(x)$ 关于 $x=a$ 的*泰勒级数*. 对任何特定的 x, 该级数可能收敛也可能发散. 若对任意特定的 x, 余项 $R_N(x)$ 当 $N \to \infty$ 时收敛于 0, 则对该 x 有

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

即在点 x 处, $f(x)$ 等于它的泰勒级数 (关于 $x=a$).

(6) 对特别的情形 $a=0$, 泰勒级数为

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n.$$

它被称为 $f(x)$ 的*麦克劳林级数*. 所以, 当看到 “麦克劳林级数” 时, 可以把它看作 “关于 $x=0$ 的泰勒级数”.

25.2 求泰勒多项式与泰勒级数

欲求特定的泰勒多项式或级数, 幸运的话, 可以通过对已知的泰勒多项式或级数运算来求得想要的多项式或级数. 我们将在 26.2 节讨论一些相应的方法. 不幸的是, 情况并不总是这样, 有时你需要从前面的总结中将 f 关于 $x=a$ 的泰勒级数分离出来:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

知道了数 a 和函数 f, 还需要求出 f 的所有导数在 $x=a$ 的值, 然后将它们代入上述公式. 然而, 这很讨厌! 求一次或两次导就已经很麻烦了, 求成百上千次导数就太荒谬了. 对于只求低次泰勒多项式来说还不是那么糟糕, 因为只需计算少量导数. 我们将在 26.2 节讨论一些可以帮你避开上面这些公式的好方法, 如果你够幸运.

另一方面, 有些函数是很容易求导的. 一个这样的例子是函数 $f(x)=\mathrm{e}^x$, 上一章我们讨论了它的麦克劳林级数. 若你不想求 f 的麦克劳林级数, 而是求它关于 $x=-2$ 的泰勒级数怎么办? 将上面公式中的 0 用 $a=-2$ 代换, 可得

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(-2)}{n!}(x+2)^n.$$

对 n 的许多值, 我们需要求 $f^{(n)}(-2)$, 所以构造一个导数表是很有帮助的. 一般地, 表的模板如下:

n	$f^{(n)}(x)$	$f^{(n)}(a)$
0		
1		
2		
3		

首先应填中间一列. 从最上一行的函数本身, 持续求导. 每次求完导后, 将结果写在表的下一行 (仍为中间一列). 当中间那列填满后, 将 $x=a$ 代入中间列的每一个值, 将相应的结果填在同行的第三列上. 注意可能要更多行, 这取决于 n 的大小或计算的快慢. 在我们的例子中, $a=-2$ 且 $f(x)$ 的所有导数均为 e^x, 所以填完的表如下:

n	$f^{(n)}(x)$	$f^{(n)}(-2)$
0	e^x	e^{-2}
1	e^x	e^{-2}
2	e^x	e^{-2}
3	e^x	e^{-2}

很清楚: 对所有 n, $f^{(n)}(-2)=\mathrm{e}^{-2}$, 若将其代入公式

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(-2)}{n!}(x+2)^n,$$

可得到 e^x 关于 $x=-2$ 的泰勒级数

$$\sum_{n=0}^{\infty}\frac{\mathrm{e}^{-2}}{n!}(x+2)^n.$$

不用求和号而将其展开是个好主意, 即

$$\mathrm{e}^{-2}+\mathrm{e}^{-2}(x+2)+\frac{\mathrm{e}^{-2}}{2!}(x+2)^2+\frac{\mathrm{e}^{-2}}{3!}(x+2)^3+\cdots.$$

还有另一个例子: 求 $\sin(x)$ 关于 $x=\pi/6$ 的泰勒级数, 写出直到第四阶的项. 我们从导数表开始.

n	$f^{(n)}(x)$	$f^{(n)}(\pi/6)$
0	$\sin(x)$	$1/2$
1	$\cos(x)$	$\sqrt{3}/2$
2	$-\sin(x)$	$-1/2$
3	$-\cos(x)$	$-\sqrt{3}/2$
4	$\sin(x)$	$1/2$

这与用来求麦克劳林级数的表类似, 不过这里是求 $\pi/6$ 处的导数而不是 0 处的导数. 写出泰勒级数的标准公式：

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n.$$

展开：

$$f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\frac{f^{(3)}(a)}{3!}(x-a)^3+\frac{f^{(4)}(a)}{4!}(x-a)^4+\cdots.$$

令 $a=\pi/6$, 将上面表中的值代入上式得到 $\sin(x)$ 关于 $x=\pi/6$ 的泰勒级数为

$$\frac{1}{2}+\frac{\sqrt{3}}{2}\left(x-\frac{\pi}{6}\right)+\frac{-1/2}{2!}\left(x-\frac{\pi}{6}\right)^2+\frac{-\sqrt{3}/2}{3!}\left(x-\frac{\pi}{6}\right)^3+\frac{1/2}{4!}\left(x-\frac{\pi}{6}\right)^4+\cdots.$$

要写出用求和号表示的形式比较难, 所以只做一个小小的化简得到：

$$\frac{1}{2}+\frac{\sqrt{3}}{2}\left(x-\frac{\pi}{6}\right)-\frac{1}{2\times 2!}\left(x-\frac{\pi}{6}\right)^2-\frac{\sqrt{3}}{2\times 3!}\left(x-\frac{\pi}{6}\right)^3+\frac{1}{2\times 4!}\left(x-\frac{\pi}{6}\right)^4+\cdots.$$

当然, 为了求四阶泰勒多项式 $P_4(x)$ (仍关于中心 $x=\pi/6$), 只需去掉后面的 "$+\cdots$". 若只想求 $P_3(x)$, 还要去掉最后那项, 则最后一项的幂次变为 3：

$$P_3(x)=\frac{1}{2}+\frac{\sqrt{3}}{2}\left(x-\frac{\pi}{6}\right)-\frac{1}{4}\left(x-\frac{\pi}{6}\right)^2-\frac{\sqrt{3}}{12}\left(x-\frac{\pi}{6}\right)^3.$$

(将 2! 换成了 2, 3! 换成了 6.) 另一方面, 若想求 $P_5(x)$, 需要在上面的表尾再加上对应于 $n=5$ 的一行, 则得到另外的项 $(x-\pi/6)^5$.

另一个例子：$(1+x)^{1/2}$ 的麦克劳林级数是什么？因为要求麦克劳林级数, 所以需要令 $a=0$. 画一个到四阶导的表：

n	$f^{(n)}(x)$	$f^{(n)}(0)$
0	$(1+x)^{1/2}$	1
1	$\frac{1}{2}(1+x)^{-1/2}$	$1/2$
2	$-\frac{1}{4}(1+x)^{-3/2}$	$-1/4$
3	$\frac{3}{8}(1+x)^{-5/2}$	$3/8$
4	$-\frac{15}{16}(1+x)^{-7/2}$	$-15/16$

现在写出麦克劳林级数的一般公式

$$f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\frac{f^{(3)}(0)}{3!}x^3+\frac{f^{(4)}(0)}{4!}x^4+\cdots,$$

将上面表中的导数值代入可得

$$1+\frac{1}{2}x+\frac{-1/4}{2!}x^2+\frac{3/8}{3!}x^3+\frac{-15/16}{4!}x^4+\cdots.$$

化简得

$$1+\frac{x}{2}-\frac{x^2}{8}+\frac{x^3}{16}-\frac{5x^4}{128}+\cdots$$

实际上, 当 x 介于 -1 和 1 之间时, 余项趋于 0(这个证明比较棘手!). 所以当 $-1<x<1$, 我们有

$$(1+x)^{1/2}=1+\frac{x}{2}-\frac{x^2}{8}+\frac{x^3}{16}-\frac{5x^4}{128}+\cdots$$

这是**二项式定理**的一个特殊情形, 即对 $-1<x<1$, 有

$$(1+x)^a=1+ax+\frac{a(a-1)}{2!}x^2+\frac{a(a-1)(a-2)}{3!}x^3+\frac{a(a-1)(a-2)(a-3)}{4!}x^4+\cdots$$

除非 a 是非负整数, 否则右边的级数在 $x>1$ 或 $x<-1$ 时都发散. (在这种情况下, 右边实际为一个多项式. 能说出为什么吗?)

25.3 用误差项估算问题

在 24.1.4 节, 我们用三阶泰勒多项式 P_3 来估算 $e^{-1/10}$, 然后用余项 R_3 来说明近似程度的好坏. 现在, 我们重新看一下这些方法并把它们一般化.

为了设置问题背景, 考虑下面两个相似的例子:

(1) 用二阶泰勒多项式估算 $e^{1/3}$, 并估算误差;

(2) 估算 $e^{1/3}$, 且误差不得大于 1/10 000.

第二个问题要比第一个难. 你也看到了, 在第一个问题中, 我们要讨论二阶泰勒多项式, 故在公式中令 $N=2$. 在第二个问题中, 我们实际上是要找到 N, 这是需要考虑的另一件事情.



用这两个问题来检验一下求解估值 (或近似) 问题的一般方法.

(1) 看一下要估算什么, 选择一个相关的函数 f. 在上面的例子中, 我们要估算 $e^{\text{某式}}$, 所以令 $f(x)=e^x$. 然后, 我们令 $x=1/3$, 这是由于 $f(1/3)=e^{1/3}$, 这就是我们要估算的量.

(2) 选一个接近 x 值的数 a, 这样 $f(a)$ 就很理想了. 这就意味着, 你应该能写出 $f(a)$ 的值, 对 $f'(a)$、$f''(a)$ 等等也一样. 在我们的例子中, 我们令 $a=0$, 因为它很接近 1/3, 且 e^0 较易计算.

(3) 如上一节所做的那样, 做 f 的导数表. 它应该有三列, 分别代表 n、$f^{(n)}(x)$ 和 $f^{(n)}(a)$ 的值. 若你知道所用的泰勒多项式的阶, 则这就是你需要的 N

的值, 一定要保证表中导数计算到第 $(N+1)$ 阶. 否则, 你就尽管一行行往下写吧, 直到厌烦为止, 只要需要, 就一直能写下去.

(4) 若你不介意估算的误差, 直接跳到第 8 步; 否则, 写出 $R_N(x)$ 的公式:

$$\boxed{R_N(x)=\frac{f^{(N+1)}(c)}{(N+1)!}(x-a)^{N+1},}$$

确保注明 "c 在 a 与 x 之间", 同时注意整个过程中用 a 的实际值替代 a.

(5) 若已知所用泰勒多项式的阶, 在上述公式中将 N 替换为该数; 若不知道, 根据你所需要的误差的大小做猜测. 误差越小, N 应该越大. 对于很多问题来说, $N=2$ 或 3 就可以了. 若该猜测值是错的, 那应该很快就能知道, 只需用较大的 N 重复这一步和下面两步.

(6) 现在, 用你想用的值代换 $R_N(x)$ 公式中的 x. 除了 c 以外, 没有其他的未知变量, 且可以用不等式写下 c 的可能范围. 在我们的例子中, 由 $a=0$ 和 $x=1/3$ 知, c 介于两者之间, 可写为 $0<c<1/3$.

(7) 求 $|R_N(x)|$ 的最大值, c 在适当的区间里. 这就是误差可能的大小. 若已知 N 的值, 就基本完成了误差估算. 若不知道, 则用你想要的误差来与实际误差比较. 若实际误差较小, 这就太好了, 你已经找到了一个较好的 N 值. 反之, 你就要回到步骤 5 再来一次. (我们将在 25.3.6 节讨论一些 $|R_N(x)|$ 极大化的方法.)

(8) 最后, 求实际的估算. 写下 $P_N(x)$ 的公式:

$$\begin{aligned}P_N(x)=f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2\\+\frac{f^{(3)}(a)}{3!}(x-a)^3+\cdots+\frac{f^{(N)}(a)}{N!}(x-a)^N.\end{aligned}$$

现在将 a 与 N 换成前面所得值而得到一个只含有 x 的公式. 最后, 写出近似

$$\boxed{f(x)\approx P_N(x),}$$

并代入所需的 x 的实际值. 右边将是你想要的量, 而左边将是近似值.

(9) 如果需要的话, 还有另一个信息: 若 $R_N(x)$ 是正的, 则估算为低估; 若 $R_N(x)$ 为负, 则估算是高估. 这些结果遵从等式

$$\boxed{f(x)=P_N(x)+R_N(x).}$$

现在, 我们来看一下有关这类问题的 5 个例子.

25.3.1　第一个例子

我们最好从前一节的两个问题开始. 在第一个问题中, 我们想用二阶泰勒多项式来估算 $e^{1/3}$, 它其实与 24.1.4 节中涉及 $e^{-1/10}$ 的问题很相似. 不管怎样, 我们还

用前面的方法. 从选择 f 开始. 因为要求幂, 令 $f(x) = \mathrm{e}^x$ 且注意 $\mathrm{e}^{1/3}$ 就是 $f(1/3)$. 最后, 我们令 $x = 1/3$, 但这还不算完, 还要选择接近 $1/3$ 的 a 使得 e^a 足够精密. 如我前面提到的, 很自然的选 0.

现在, 该填导数表了：

n	$f^{(n)}(x)$	$f^{(n)}(0)$
0	e^x	1
1	e^x	1
2	e^x	1
3	e^x	1

我求到 3 阶导, 因为它刚好大于 2, 我们需要二阶泰勒多项式 (即 $N = 2$). 好, 继续. 误差项为

$$R_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!}x^{N+1},$$

其中 c 介于 0 与 x 之间. 注意在 $R_N(x)$ 的标准公式中, 我将 a 换成了 0. 现在, 我们知道 $N = 2$, 所以实际需要

$$R_2(x) = \frac{f^{(3)}(c)}{3!}x^3 = \frac{\mathrm{e}^c}{6}x^3.$$

在前面的表中, 将中间一列的最后一行的 x 换成 c, 得到 $f^{(3)}(c) = \mathrm{e}^c$. 现在将 x 换为 $1/3$ 可得

$$R_2(1/3) = \frac{\mathrm{e}^c}{6}(1/3)^3 = \frac{\mathrm{e}^c}{162};$$

这里 c 介于 0 与 $x = 1/3$ 之间, 故 $0 < c < 1/3$. 取绝对值有：

$$|R_2(1/3)| = \left|\frac{\mathrm{e}^c}{162}\right| = \frac{\mathrm{e}^c}{162},$$

因为 e^c 必为正. 接下来, 我们需要最大化 $|R_2(1/3)|$. 由于 e^c 关于 c 递增, 最大值出现在 $c = 1/3$ 时. 这就有

$$|R_2(1/3)| = \frac{\mathrm{e}^c}{162} < \frac{\mathrm{e}^{1/3}}{162}.$$

似乎有一个问题, 我们不知道 $\mathrm{e}^{1/3}$ 是多少. 这其实是该问题的关键点! 没关系, 我们粗略高估一下 $\mathrm{e}^{1/3}$. 你知道, $\mathrm{e} < 8$, 所以 $\mathrm{e}^{1/3} < 8^{1/3}$, 而 $8^{1/3}$ 为 2. 我为什么选 8 呢? 因为我可以什么都不用想就直接取它的三次方根! 总之, 运用不等式 $\mathrm{e}^{1/3} < 2$, 前面 $|R_2(1/3)|$ 的不等式变为

$$|R_2(1/3)| = \frac{\mathrm{e}^c}{162} < \frac{\mathrm{e}^{1/3}}{162} < \frac{2}{162} = \frac{1}{81}.$$

所以误差不大于 $1/81$. 我们仍需求估算值. 写下 $P_2(x)$ 的公式, 并令 $a = 0$：

$$P_2(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2.$$

根据前面的表, 将 $f(0)$、$f'(0)$ 和 $f''(0)$ 换为 1：

$$P_2(x) = 1 + x + \frac{1}{2}x^2.$$

最后, 令 $x = 1/3$ 可得

$$P_2(1/3) = 1 + \frac{1}{3} + \frac{1}{2}\left(\frac{1}{3}\right)^2 = \frac{25}{18}.$$

由于 $f(x) \approx P_2(x)$, 我们有

$$f(1/3) \approx P_2(1/3).$$

根据 $f(x) = \mathrm{e}^x$, 我们有

$$\mathrm{e}^{1/3} = f(1/3) \approx P_2(1/3) = \frac{25}{18}.$$

我们已经得到了 $|R_2(1/3)| < 1/81$, 所以估算值至少精确到 1/81. 其实, 因为 $R_2(1/3)$ 是正的, 所以估算值 25/18 相对于 $\mathrm{e}^{1/3}$ 真实值是低估了.

25.3.2　第二个例子

我们将讨论 25.3 节的第二个例子：估算 $\mathrm{e}^{1/3}$ 的值, 且误差小于 1/10 000. 与前一个例子一样, 我们令 $f(x) = \mathrm{e}^x, a = 0$, 最后令 $x = 1/3$, 我们有

$$R_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!}x^{N+1},$$

其中 c 介于 0 与 x 之间. 我们已经从前一个例子知道, 不能令 $N = 2$, 因为此时会得到一个最大的误差 1/81, 而我们需要误差小于 1/10 000. 所以, 来看一下 $N = 3$ 是否可行. 现在误差项为

$$R_3(x) = \frac{f^{(4)}(c)}{4!}x^4 = \frac{\mathrm{e}^c}{24}x^4,$$

其中 c 介于 0 与 x 之间. 令 $x = 1/3$ 可得

$$R_3(1/3) = \frac{\mathrm{e}^c}{24}\left(\frac{1}{3}\right)^4 = \frac{\mathrm{e}^c}{24 \times 81},$$

其中 $0 < c < 1/3$. 我们引用前一节的结论, 当 c 介于 0 与 1/3 之间时, $\mathrm{e}^c < 2$：

$$|R_3(1/3)| = \frac{|\mathrm{e}^c|}{24 \times 81} < \frac{2}{24 \times 81} = \frac{1}{972}.$$

这个结果并不小于 1/10 000, 所以 $N = 3$ 不够大. 再试一下 $N = 4$. 重复上面的步骤, 有

$$R_4(x) = \frac{f^{(5)}(c)}{5!}x^5 = \frac{\mathrm{e}^c}{120}x^5,$$

所以令 $x = 1/3$, 可知

$$R_4(1/3) = \frac{\mathrm{e}^c}{120}\left(\frac{1}{3}\right)^5 = \frac{\mathrm{e}^c}{120 \times 243}.$$

c 还是介于 0 与 1/3 之间, 同样有 $\mathrm{e}^c < 2$, 所以

$$|R_4(1/3)| < \frac{2}{120 \times 243} = \frac{1}{14\,580}.$$

(别急着用计算器来计算最后的分数, 再想一下, 其实可以将 2/120 化简为 1/60, 然后算出 6×243, 再乘以 10, 最后写在分母上.) 不管怎样, 我们知道 $|R_4(1/3)|$ 远小于 1/10 000, 所以目的达到了: 令 $N = 4$, 那估算值是多少呢? 我们需要求出 $P_4(1/3)$. 一般地, 当 $a = 0$, 四阶泰勒多项式 P_4 为

$$P_4(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!},$$

所以

$$P_4\left(\frac{1}{3}\right) = 1 + \frac{1}{3} + \frac{(1/3)^2}{2} + \frac{(1/3)^3}{6} + \frac{(1/3)^4}{24} = 1 + \frac{1}{3} + \frac{1}{18} + \frac{1}{162} + \frac{1}{1944} = \frac{2713}{1944},$$

即

$$\mathrm{e}^{1/3} = f(1/3) \approx P_4(1/3) = \frac{2713}{1944}.$$

所以, 我们可以将前一个例子中的估算值 25/18 替换为更好的估算, 即 2713/1944. 这个新的估算值保证与 $\mathrm{e}^{1/3}$ 真实值的误差在 1/10 000 以内. 验证一下, 我用计算器算出 2713/1944 精确到 5 位小数的值为 1.395 58, 而 $\mathrm{e}^{1/3}$ 精确到 5 位小数的值为 1.395 61, 最多相差 0.000 04, 显然在允许的范围 $1/10\,000 = 0.0001$ 之内.

25.3.3 第三个例子

这里有一个问题: 估算 $\sqrt{27}$, 误差不大于 1/250. 根据前面的方法, 我们需要选择一个合适的函数 f 以及 a 和 x 的值. 较好的选择是令 $f(x) = \sqrt{x}$, 或者 $f(x) = x^{1/2}$, 随便哪个都行. 我们要估算 $f(27) = \sqrt{27}$ 的值, 所以令 $x = 27$. 现在要找一个接近 27 且易求平方根的数. 似乎 25 就可以, 我们令 $a = 25$, 这是第一步. 第二步是填一个导数表:



n	$f^{(n)}(x)$	$f^{(n)}(25)$
0	$x^{1/2}$	5
1	$\frac{1}{2}x^{-1/2}$	1/10
2	$-\frac{1}{4}x^{-3/2}$	$-1/500$
3	$\frac{3}{8}x^{-5/2}$	$3/8 \times 1/5^5$

记住, 要填这个表, 在首行的中间一列填上 $x^{1/2}$, 然后连续求几次导, 将结果填在中间列的后面几行. 最后, 右边是将值 $a = 25$ 代入所得的值. 困难的是, 我们不知道这个表需要填到第几行. 可能需要更多行.

现在我们来看误差项

$$R_N(x)=\frac{f^{(N+1)}(c)}{(N+1)!}(x-25)^{N+1},$$

其中 c 介于 x 与 25 之间. 我们关注的是 $x=27$, 将其代入得到

$$R_N(27)=\frac{f^{(N+1)}(c)}{(N+1)!}(27-25)^{N+1}=\frac{f^{(N+1)}(c)}{(N+1)!}2^{N+1},$$

其中 $25\leqslant c\leqslant 27$. 感到幸运了吗? 或许 $N=0$ 就够好了! 我们来试一下:

$$|R_0(27)|=\left|\frac{f'(c)}{1!}(27-25)^1\right|=\frac{1}{2}c^{-1/2}\times 2=c^{-1/2},$$

其中我们利用前面的表来求 $f'(c)$ 并去掉了绝对值, 因为所有数都是正的. 现在的大问题是, 对给定的 $25\leqslant c\leqslant 27$, $c^{-1/2}$ 有多大? 注意 $c^{-1/2}$ 关于 c 递减, 所以当 $c=25$ 时有最大值. 因为 $c^{-1/2}$ 为 $25^{-1/2}=1/5$, 所以, 有

$$|R_0(27)|=c^{-1/2}\leqslant 1/5.$$

故误差可能高达 1/5. 有点太高了, 我们需要误差不大于 1/250. 所以选择 $N=0$ 显然有点太过于乐观了! 我们需要更好点. 试一下 $N=1$, 则

$$|R_1(27)|=\left|\frac{f''(c)}{2!}(27-25)^2\right|=\left|-\frac{1}{4}c^{-3/2}\times\frac{1}{2!}\times 2^2\right|=\frac{c^{-3/2}}{2}.$$

同样用前面的表来求 $f''(c)$. 这次我要用绝对值, 因为 $R_1(27)$ 是负的 (是的, 有些估高了). 还是当 c 最小时 $c^{-3/2}$ 最大, 即 $c=25$, 这时表达式为 $25^{-3/2}=1/125$, 所以

$$|R_1(27)|=\frac{c^{-3/2}}{2}\leqslant\frac{1}{125}\times\frac{1}{2}=\frac{1}{250}.$$

这就意味着误差不大于 1/250, 正是我们想要的. 因此取 $N=1$, 我们只需求 $P_1(27)$. (因为 $N=1$, 这里我们其实运用了线性化.) 总之, 我们知道了

$$P_1(x)=f(25)+f'(25)(x-25)=5+\frac{1}{10}(x-25),$$

其中 $f(25)$ 和 $f'(25)$ 的值可从前面的表中得到, 令 $x=27$, 有

$$P_1(27)=5+\frac{1}{10}(27-25)=\frac{26}{5}.$$

我们得到 $\sqrt{27}$ 近似等于 26/5 的结论, 这两个数之间的差在 1/250 之内, 且 26/5 高于 $\sqrt{27}$ (因为误差项 $R_1(27)$ 是负的). 事实上, 计算器算出的 $\sqrt{27}$ 约为 5.196 15, 与 26/5 = 5.2 的差在 1/250 之内. 对 $N=2$ 或更大值的情况, 估算值不会错, 反而会更好, 只不过数会更不整洁.

25.3.4　第四个例子



为了提出本节的问题, 我们将前面的问题做个小的变动. 我们将 $\sqrt{27}$ 换成 $\sqrt{23}$, 欲估算 $\sqrt{23}$ 的误差不大于 1/250 的值. 这没比前面的例子难多少, 是吧? 然而, 也

不尽然, 我们来看看. 我们仍将采用 $f(x) = x^{1/2}$, $a = 25$ 的泰勒级数, 不过这里需要将 $x = 27$ 换为 $x = 23$. 我们来看一下余项 R_1:

$$|R_1(23)| = \left|\frac{f''(c)}{2!}(23-25)^2\right| = \left|-\frac{1}{4}c^{-3/2} \times \frac{1}{2!} \times (-2)^2\right| = \frac{c^{-3/2}}{2}.$$

这就是误差项! 不过现在与前一个例子有个很重要的不同: c 介于 23 和 25 之间. 所以 $\frac{1}{2}c^{-3/2}$ 有多大呢? 这个量仍关于 c 递减, 所以随着 c 的减小, 其值变成最大值, 即当 $c = 23$ 时值最大. 因此有如下的估算:

$$|R_1(23)| = \frac{c^{-3/2}}{2} \leqslant \frac{23^{-3/2}}{2}.$$

不幸的是, $23^{-3/2}$ 并不比 $25^{-3/2}$ 好算. 我们唯一可以肯定的是这种情况不够好. 你知道, $\frac{1}{2} \cdot 25^{-3/2} = 1/250$, 但 $\frac{1}{2} \cdot 23^{-3/2}$ 大于 $1/250$, 所以太大了. 所以 $N = 1$ 不行, 需要试一下 $N = 2$.

取 $N = 2$ 并运用 25.3.3 节的表, 有

$$|R_2(23)| = \left|\frac{f^{(3)}(c)}{3!}(23-25)^3\right| = \left|-\frac{3}{8}c^{-5/2} \times \frac{1}{3!} \times (-2)^3\right| = \frac{c^{-5/2}}{2},$$

其中 $23 \leqslant c \leqslant 25$. 这次当 $c = 23$ 时, $c^{-5/2}$ 还是最大的, 因此有

$$|R_2(23)| = \frac{c^{-5/2}}{2} \leqslant \frac{23^{-5/2}}{2}.$$

这个够好吗? 没有计算器, 我们不得不寻找一些估算 $23^{-5/2}$ 的方法. 朋友, 你怎么来实现呢? 我能想到的最好办法就是找一个小于 23 的数, 并且这个数的 $-5/2$ 次幂是容易算出来的. 那应该是 16, 而 $16^{-5/2} = 1/4^5 = 1/1024$, 所以

$$|R_2(23)| \leqslant \frac{23^{-5/2}}{2} \leqslant \frac{16^{-5/2}}{2} = \frac{1}{1024} \times \frac{1}{2} = \frac{1}{2048}.$$

这个值当然小于 $1/250$, 所以采用 $N = 2$ 是可以的, 我们就可以用 $P_2(23)$ 了. 现在

$$\begin{aligned} P_2(x) &= f(25) + f'(25)(x-25) + \frac{f''(25)}{2!}(x-25)! \\ &= 5 + \frac{1}{10}(x-25) - \frac{1}{500 \times 2}(x-25)^2. \end{aligned}$$

(再一次利用那个表), 将 x 用 23 代换, 我们有

$$P_2(23) = 5 + \frac{1}{10}(23-25) - \frac{1}{1000}(23-25)^2 = 5 - \frac{2}{10} - \frac{4}{1000} = \frac{1199}{250}.$$

因此对 $\sqrt{23}$ 的估算值是 $1199/250$. 用计算器得出最后分数的结果等于 4.796, 而 $\sqrt{23}$ 的计算结果为 4.795 83. 这两个数的差的确在 $1/250$ 范围内.

25.3.5 第五个例子

我们再来看一个例子: 用三阶泰勒级数估算 $\cos(\pi/3 - 0.01)$ 的值, 并给出该估



算的精确度. 我们需要选择一个函数, 显而易见的函数是 $f(x)=\cos(x)$, 所以我们要令 $x=\pi/3-0.01$. 那余弦值易求且接近于 x 的数是什么呢? 显然 $a=\pi/3$ 是一个自然的候选项. 故我们得到如下的表:

n	$f^{(n)}(x)$	$f^{(n)}(\pi/3)$
0	$\cos(x)$	$1/2$
1	$-\sin(x)$	$-\sqrt{3}/2$
2	$-\cos(x)$	$-1/2$
3	$\sin(x)$	$\sqrt{3}/2$
4	$\cos(x)$	不需要

误差项 $R_3(x)$ 为

$$R_3(x)=\frac{f^{(4)}(c)}{4!}\left(x-\frac{\pi}{3}\right)^4=\frac{\cos(c)}{24}\left(x-\frac{\pi}{3}\right)^4,$$

其中 c 介于 x 与 $\pi/3$ 之间. 注意, 我们需要的是 $f^{(4)}(c)$ 而不是 $f^{(4)}(\pi/3)$, 这就解释了表中出现的 “不需要”. 当 $x=\pi/3-0.01$, 我们有

$$R_3\left(\frac{\pi}{3}-0.01\right)=\frac{\cos(c)}{24}\left(\frac{\pi}{3}-0.01-\frac{\pi}{3}\right)^4=\frac{\cos(c)}{24}(-0.01)^4=\frac{\cos(c)}{24\times10^8}.$$

(这里我们使用了 $(-0.01)^4=(0.01)^4=(10^{-2})^4=10^{-8}$.) 现在只需估算误差项的绝对值. 鉴于 $|\cos(c)|\leqslant1$, 我们有

$$\left|R_3\left(\frac{\pi}{3}-0.01\right)\right|=\frac{|\cos(c)|}{24\times10^8}\leqslant\frac{1}{24\times10^8}=\frac{1}{2\,400\,000\,000}.$$

太好了, 我们知道运用 $P_3(\pi/3-0.01)$ 来估算 $\cos(\pi/3-0.01)$ 会使得估算值精确到很小的数 $1/2\,400\,000\,000$. 那 $P_3(\pi/3-0.01)$ 是多少呢? 根据公式有

$$P_3(x)=f\left(\frac{\pi}{3}\right)+f'\left(\frac{\pi}{3}\right)\left(x-\frac{\pi}{3}\right)+\frac{1}{2!}f''\left(\frac{\pi}{3}\right)\left(x-\frac{\pi}{3}\right)^2+\frac{1}{3!}f^{(3)}\left(\frac{\pi}{3}\right)\left(x-\frac{\pi}{3}\right)^3.$$

应用上面的导数表, 其变为

$$P_3(x)=\frac{1}{2}-\frac{\sqrt{3}}{2}\left(x-\frac{\pi}{3}\right)-\frac{1}{2}\times\frac{1}{2}\left(x-\frac{\pi}{3}\right)^2+\frac{1}{6}\times\frac{\sqrt{3}}{2}\left(x-\frac{\pi}{3}\right)^3.$$

令 $x=\pi/3-0.01$ 并化简, 结果是

$$\begin{aligned}P_3\left(\frac{\pi}{3}-0.01\right)&=\frac{1}{2}-\frac{\sqrt{3}}{2}(-0.01)-\frac{1}{4}(-0.01)^2+\frac{\sqrt{3}}{12}(-0.01)^3\\&=\frac{1}{2}+\frac{\sqrt{3}}{200}-\frac{1}{40\,000}-\frac{\sqrt{3}}{12\,000\,000}.\end{aligned}$$

这个表达式看起来很麻烦, 但其实还不错, 唯一棘手的量是 $\sqrt{3}$, 不过它本身是容易估算的, 至少表达式中没有三角函数了. 总之, 由于 $f(\pi/3-0.01)$ 近似等于 $P_3(\pi/3-0.01)$, 我们有

$$\cos\left(\frac{\pi}{3}-0.01\right)=f\left(\frac{\pi}{3}-0.01\right)\approx\frac{1}{2}+\frac{\sqrt{3}}{200}-\frac{1}{40\,000}-\frac{\sqrt{3}}{12\,000\,000},$$

精确到 1/2 400 000 000 之内.

25.3.6 误差项估算的一般方法

在前面所有例子中, 我们都要对在某区间内取值的 c 来估算 $\left|f^{(N+1)}(c)\right|$. 我们总结下一般的对策.

(1) 不管 c 是多少, 你总能使用标准的不等式 $|\sin(c)| \leqslant 1$ 和 $|\cos(c)| \leqslant 1$.

(2) 若函数 $f^{(N+1)}$ 是递增的, 则它的值在右端点最大. 在前两个例子中, 我们要求 e^c 的最大值, 其中 $0 < c < 1/3$. 由于 e^c 关于 c 递增, 所以可以说 $\mathrm{e}^c < \mathrm{e}^{1/3}$. 另一方面, 在 24.1.4 节的例子中, 我们也需要最大化 e^c, 不过那次 $-1/10 < c < 0$. 同样, 由于 e^c 关于 c 递增, 因而这个最大值就是 $\mathrm{e}^0 = 1$, 即 $\mathrm{e}^c < \mathrm{e}^0 = 1$.

(3) 若函数 $f^{(N+1)}$ 是递减的, 则它的最大值 $f^{(N+1)}(c)$ 出现在区间的左端点. 例如, 若已知 c 介于 1 和 5 之间, 则最大值 $1/(3+c)^4$ 出现在区间 $[1,5]$ 的左端点, 因为 $1/(3+c)^4$ 关于 c 递减. 所以上面的表达式在 $c=1$ 时最大, 相应的值为 $1/4^4 = 1/256$.

(4) 一般地, 为了求最大值, 可能还要求函数 $f^{(N+1)}$ 的临界点. (具体求法见 11.1.1 节.)

25.4 误差估算的另一种方法

回想一下交错级数判别法 (见 22.5.4 节). 该判别法表明若级数是交错的, 且各项的绝对值递减趋于 0, 则级数收敛. 收敛的原因是, 部分和与真实极限值之间就如儿童摇摇乐车：这个部分和大点, 下一个部分和小点, 再下一个部分和大点, 等等. 每次, 部分和都更接近真实极限值, 就像摇摇乐车正在失去动力. 方法就是在级数中的每个点, 每加一项都超越真实值, 所以整个误差小于下一项的绝对值.

我们用符号来表述. 假设从某函数 f 开始, 求它关于 $x=a$ 的泰勒级数. 若碰巧你还知道级数对某些特定的 x 值收敛于 $f(x)$(就像我们讨论的一些函数一样), 则可以写为

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

对那些你感兴趣的特定的 x 值, 上述级数若是各项绝对值递减趋于 0 的交错级数, 则误差小于下一项. 即

$$|R_N(x)| \leqslant \left|\frac{f^{(N+1)}(a)}{(N+1)!}(x-a)^{N+1}\right|.$$

这里没有讨厌的 c, 这是足以成为我们运用这个理想结论的原因. 记住, 上述结论只有当级数满足交错级数的三个条件时才成立!



下面是该方法适用的例子. 假设我们欲用麦克劳林级数来求定积分

$$\int_0^1 \frac{1-\cos(t)}{t^2}\mathrm{d}t,$$

误差不大于 1/3000 的估算值. 该积分好像是一个瑕点在 $t=0$ 的反常积分, 但其实 $t=0$ 不是瑕点. 由洛必达法则可知

$$\lim_{t\to 0}\frac{1-\cos(t)}{t^2} \overset{\text{l'H}}{=} \lim_{t\to 0}\frac{\sin(t)}{2t} = \frac{1}{2}.$$

即, 被积函数在 $t=0$ 并没有趋于无穷, 所以积分不是反常的. 不管怎样, 刚刚只是观察, 现在我们要解决问题.

第一个有用的方法是先构造一个像上述积分的函数, 令

$$f(x) = \int_0^x \frac{1-\cos(t)}{t^2}\mathrm{d}t.$$

则我们要估算的积分是 $f(1)$. 我们要求 f 的麦克劳林级数. 为此, 将 $\cos(t)$ 用它的麦克劳林级数代换, 该级数已在 24.2.3 节求得, 即

$$f(x) = \int_0^x \frac{1-\left(1-\frac{t^2}{2!}+\frac{t^4}{4!}-\frac{t^6}{6!}+\frac{t^8}{8!}-\cdots\right)}{t^2}\mathrm{d}t.$$

若稍作化简, 可写成

$$f(x) = \int_0^x \left(\frac{1}{2!}-\frac{t^2}{4!}+\frac{t^4}{6!}-\frac{t^6}{8!}+\cdots\right)\mathrm{d}t.$$

现在求积分并计算在端点处的值：

$$\begin{aligned} f(x) &= \left(\frac{t}{2!}-\frac{t^3}{3\times 4!}+\frac{t^5}{5\times 6!}-\frac{t^7}{7\times 8!}+\cdots\right)\Bigg|_0^x \\ &= \frac{x}{2!}-\frac{x^3}{3\times 4!}+\frac{x^5}{5\times 6!}-\frac{x^7}{7\times 8!}+\cdots. \end{aligned}$$



尝试将上式用求和号表示是一个很好的做法. 总之, 现在可将 $x=1$ 代入得

$$f(1) = \int_0^1 \frac{1-\cos(t)}{t^2}\mathrm{d}t = \frac{1}{2!}-\frac{1}{3\times 4!}+\frac{1}{5\times 6!}-\frac{1}{7\times 8!}+\cdots.$$

说实话, 这里我将两个更快的方法放在了一起. 首先, 我将 $\cos(t)$ 用它的麦克劳林级数代替. 还好我们已经在 24.2.3 节知道这对所有 t 都成立. 其次, 我对无穷级数逐项求积分, 并声明对所有 x 都可以这么做. 我们将在 26.2.3 节看到这么做是可以的 (虽然我们不会对其证明). 总之, 上面的等式是正确的. 现在给定的积分有一个无穷级数的表达式.

现在唯一的问题是, 要求与真实值误差在 1/3000 内的近似值需取多少项？注意该级数是各项递减趋于 0 的交错级数, 那么我们可以运用下一项的绝对值大于误差的结论. 例如, 若用首项 $1/2!$ 近似积分, 则误差不大于 $1/(3\times 4!)$, 即 $1/72$. 这

也太大了. 那用前两项来近似该积分怎么样? 即,

$$\int_0^1 \frac{1-\cos(t)}{t^2}\mathrm{d}t \approx \frac{1}{2!} - \frac{1}{3\times 4!} = \frac{35}{72}$$

怎样? 那么误差小于下一项的绝对值:

$$|\text{误差}| \leqslant \frac{1}{5\times 6!} = \frac{1}{5\times 720} = \frac{1}{3600}.$$

这小于我们的容忍度 1/3000, 很好. 我们完全可以说积分近似等于 35/72, 误差小于 1/3000. (我们甚至可以说 35/72 是低估的, 为什么?) 我用处理这类问题的计算机程序求了一下积分, 得到积分值约为 0.486 385, 而计算器计算的 35/72 值等于 0.486 111 (精确到 6 位小数), 这两个数的差的确在 1/3000 内.



作为练习, 试着用与上面相同的方法近似

$$\int_0^{1/2} \frac{\sin(t)}{t}\mathrm{d}t,$$

误差为 1/1000. (你会用到 $\sin(t)$ 的麦克劳林级数, 这个可在 26.2 节找到.)

第 26 章　泰勒级数和幂级数：如何解题

本章, 我们将讨论涉及泰勒级数、泰勒多项式和幂级数的四类不同问题：

- 如何确定幂级数收敛或发散的区间;
- 如何利用现有泰勒级数来求其他的泰勒级数和泰勒多项式;
- 利用泰勒级数或泰勒多项式求导;
- 利用麦克劳林级数求极限.

26.1　幂级数的收敛性

假定有一个关于 $x = a$ 的幂级数

$$\sum_{n=0}^{\infty} a_n(x-a)^n.$$

由几何级数的例子可见, 一个幂级数可能对某些 x 收敛, 而对某些 x 发散. 我们想问的一个问题是：对上面给定的幂级数, x 取何值时收敛, 取何值时发散？另外, 假设级数对某特定的 x 收敛, 若能确定该收敛是绝对收敛还是条件收敛就好了. 所以, 我们来看一下可能会发生什么, 然后好好利用这些观察所得结果.

26.1.1　收敛半径

我们想知道什么样的 x 能使幂级数 $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ 收敛. 表面上看, 我们似乎必须回答无穷多个问题, 因为有无穷多个 x 的值需要代入并验证级数收敛与否. 我们画一个数轴来表示 x 的不同值. 对每个使级数收敛的 x, 都在它上面打个对号; 而对使级数发散的 x 就打个叉号. (当然, 我们不是对每个 x 都这么做, 若如此, 图就太挤了! 只标一部分, 得到结论就可以了.) 例如, 几何级数 $\sum\limits_{n=0}^{\infty} x^n$ 当 $-1 < x < 1$ 时收敛, 其他情况均发散, 如图 26-1 所示.

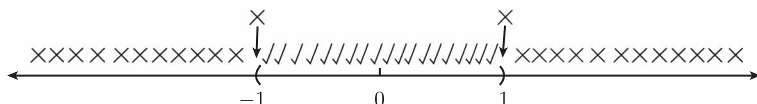


图　26-1

注意我对在端点 -1 和 1 处的发散做了特别标注.

另外, 我们已经知道级数

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}$$

对所有 x 收敛 (当然, 收敛到 e^x), 如图 26-2 所示.

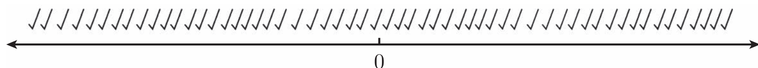


图 26-2

看起来似乎是难以预测的. 我们可以确定的是, 幂级数在 $x=a$ 处都收敛. 其实, 若将 $x=a$ 代入

$$\sum_{n=0}^{\infty}a_n(x-a)^n = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots,$$

就能知道除了 a_0 外, 其他项都没有了. 因此, 级数显然收敛 (到 a_0). 不幸的是, $x=a$ 是我们唯一可以确定收敛性的值. 那其他的值呢? 可能会是对号和叉号的大杂烩, 如图 26-3 所示.

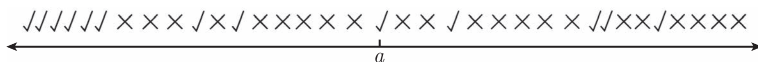


图 26-3

事实证明, 幂级数是不会像上图这样. 具体来说, 只可能出现如下三种可能性.

(1) 存在某数 $R>0$, 被称为幂级数的收敛半径, 如图 26-4 所示.

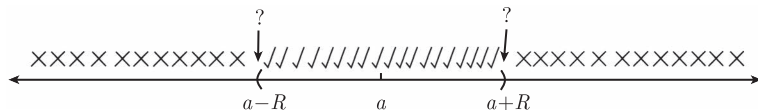


图 26-4

该图的解释如下.

- 幂级数在区域 $|x-a|<R$ 内收敛 (也可将该条件写为 $a-R<x<a+R$), 所以图像在那个区间是对号.
- 幂级数在区域 $|x-a|>R$ 内发散 (也可将该条件写为 $x<a-R$ 或 $x>a+R$), 所以图像在那个区间是叉号.
- 在两个特殊点 $|x-a|=R$ (即 $x=a+R$ 和 $x=a-R$) 处, 幂级数可能绝对收敛、条件收敛或发散. 这需要分别对这两个点进行讨论, 所以上图在这两个点处是问号. 我将称这样的点称为 “端点”.

(2) 幂级数可能对所有的 x 均绝对收敛, 这种情况下的图像如图 26-5 所示.

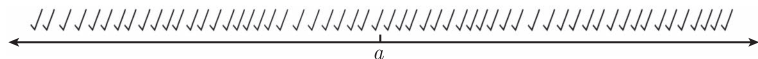


图 26-5

在这种情况下, 我们说收敛半径为 ∞. 如我们前面所见, 这样的一个例子是 e^x 的幂级数

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

其他的例子包含 $\sin(x)$ 和 $\cos(x)$ 的麦克劳林级数.

(3) 幂级数可能只在 $x=a$ 时收敛, 而对其他所有的 x 均发散. 在这种情况下, 收敛半径为 0, 我们很快就会知道, 级数

$$\sum_{n=0}^{\infty}n!x^n$$

就是这种情形. 这种情况的图像如图 26-6 所示.

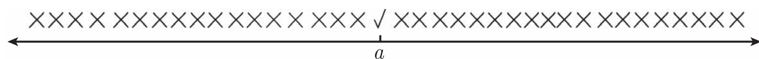


图　26-6

当然, 我还没有说为什么这些是仅有的可能. 不过很快就可以弄清楚了!

26.1.2　求收敛半径和收敛区域



给定一个幂级数, 如何求收敛半径? 答案是用比式判别法. 有时, 根式判别法会更有效, 但比式判别法对大多数问题更合适. (比式判别法和根式判别法的更多细节分别见 23.3 节和 23.4 节.) 这里是一般的方法.

(1) 写出比值绝对值的极限, 常常为

$$\lim_{n\to\infty}\left|\frac{a_{n+1}(x-a)^{n+1}}{a_n(x-a)^n}\right|=\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right||x-a|.$$

若使用的是根式判别法, 则得到

$$\lim_{n\to\infty}|a_n(x-a)^n|^{1/n}=\lim_{n\to\infty}|a_n|^{1/n}|x-a|.$$

(2) 算出极限. 注意, 极限是在 $n\to\infty$ 时而不是 $x\to\infty$ 时. 它们的差别很大! 无论是运用比式判别法还是根式判别法, 答案都形如 $L|x-a|$, 其中 L 可能是一个有限值、0 或者 ∞. 重要的是结果中有因子 $|x-a|$.

(3) 不管是比式判别法还是根式判别法, 重要的是极限 $L|x-a|$ 是小于 1, 大于 1, 还是等于 1. 所以, 若 L 是正的, 则除以 L 就能知道一切：若 $|x-a|<1/L$, 则幂级数绝对收敛; 若 $|x-a|>1/L$, 则幂级数发散; 若 $|x-a|=1/L$, 则得不到结论, 需要讨论两个端点. 这是前一节的第一种情形, 收敛半径是 $1/L$.

(4) 若 $L=0$, 则不论 x 取何值, 比式的极限都为 0. 由于 $0<1$, 这意味着幂级数对所有的 x 值都绝对收敛, 所以, 这是前一节的第二种情形, 收敛半径为 ∞.

(5) 若 $L=\infty$, 则看起来似乎幂级数永不收敛. 其实, 当 $x=a$ 时幂级数一定收敛, 但幂级数对其他的任何 x 值都发散. 所以, 这里前一节的第三种情形, 收敛半

径为 0.

这或多或少地说明了我们为什么必然得到前一节的三种情形之一. 不过, 这些仍很抽象, 还需要用一系列的例子来加以说明.



首先, 考虑幂级数

$$\sum_{n=2}^{\infty}\frac{x^n}{n\ln(n)}.$$

我们采用比式判别法. 我们从取通项 $x^n/(n\ln(n))$ 开始, 并把它作为一个大分数的分母; 然后选取大分数的分子, 还是从通项 $x^n/(n\ln(n))$ 开始, 不过这次将每个 n 用 $n+1$ 代换; 最后, 取绝对值, 然后取 $n\to\infty$ 的极限. 所以, 我们需要考虑的是

$$\lim_{n\to\infty}\left|\frac{\dfrac{x^{n+1}}{(n+1)\ln(n+1)}}{\dfrac{x^n}{n\ln(n)}}\right|.$$

这与普通的用比式判别法的级数问题一样：只需合并同类项. 可得

$$\begin{aligned}\lim_{n\to\infty}\left|\frac{\dfrac{x^{n+1}}{(n+1)\ln(n+1)}}{\dfrac{x^n}{n\ln(n)}}\right| &= \lim_{n\to\infty}\left|\frac{x^{n+1}}{x^n}\frac{n}{n+1}\frac{\ln(n)}{\ln(n+1)}\right|\\ &= \lim_{n\to\infty}|x|\frac{n}{n+1}\frac{\ln(n)}{\ln(n+1)} = |x|.\end{aligned}$$



同样, 极限是在 $n\to\infty$ 时, 这就是将 $n/(n+1)$ 和 $\ln(n)/\ln(n+1)$ 换成 1 的原因. (对对数运用洛必达法则, 细节自行完成.) 总之, 比式的极限为 $|x|$, 故由比式判别法, 我们的幂级数当 $|x|<1$ 时绝对收敛, 当 $|x|>1$ 时发散, 即收敛半径为 1. 我们仍需讨论 $x=1$ 和 $x=-1$ 时的情形. 先看 $x=1$, 将 $x=1$ 代入, 则原幂级数变为

$$\sum_{n=2}^{\infty}\frac{1^n}{n\ln(n)}=\sum_{n=2}^{\infty}\frac{1}{n\ln(n)}.$$

它收敛吗？你可运用积分判别法得到, 它发散 (或见 23.5 节). 现在将 $x=-1$ 代入原幂级数可得

$$\sum_{n=2}^{\infty}\frac{(-1)^n}{n\ln(n)}.$$



它不绝对收敛, 事实上, 将该级数的各项用它们的绝对值代换就是当 $x=1$ 时的级数, 这个级数刚刚已被证得是发散的. 另一方面, 上面 $x=-1$ 对应的级数可由交错级数判别法证得是收敛的 (用 23.7 节的方法, 可自行写出具体细节). 于是, 我们知道在点 $x=-1$ 处条件收敛. 总之, 幂级数在 $-1<x<1$ 时绝对收敛, 当 $x=-1$ 时条件收敛, 对其他的所有 x 都发散. 图像如图 26-7 所示.

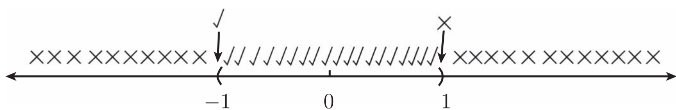


图　26-7

现在考虑

$$\sum_{n=2}^{\infty} \frac{x^n}{n(\ln(n))^2}.$$

这与前一个问题几乎一样, 我们来看看. 我们有

$$\lim_{n\to\infty} \left| \frac{\dfrac{x^{n+1}}{(n+1)(\ln(n+1))^2}}{\dfrac{x^n}{n(\ln(n))^2}} \right| = \lim_{n\to\infty} \left| \frac{x^{n+1}}{x^n} \frac{n}{n+1} \frac{(\ln(n))^2}{(\ln(n+1))^2} \right|$$

$$= \lim_{n\to\infty} |x| \frac{n}{n+1} \left(\frac{\ln(n)}{\ln(n+1)} \right)^2,$$

它仍然可以化简到 $|x|$. 故幂级数还是在 $|x| < 1$ 时绝对收敛, 在 $|x| > 1$ 时发散. 因此, 收敛半径是 1. 对于端点, 我们令 $x = 1$:

$$\sum_{n=2}^{\infty} \frac{1^n}{n(\ln(n))^2} = \sum_{n=2}^{\infty} \frac{1}{n(\ln(n))^2}.$$

如 23.5 节所述, 你可运用积分判别法得到该级数收敛, 由于各项均为正, 所以收敛为绝对收敛. 现在, 代入 $x = -1$, 我们得到

$$\sum_{n=2}^{\infty} \frac{(-1)^n}{n(\ln(n))^2}.$$

各项取绝对值对应的级数为

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln(n))^2},$$

这与 $x = 1$ 时的级数一样, 所以它绝对收敛. 我们得到结论：当 $-1 \leqslant x \leqslant 1$ 时幂级数绝对收敛, 且级数对其他所有 x 发散, 如图 26-8 所示.

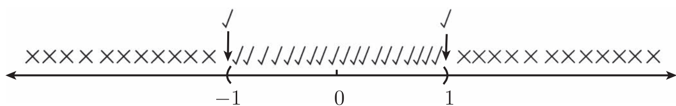


图　26-8

所以, 除了在端点 1 和 −1 处不同之外, 它与前一个例子一样.

那级数

$$\sum_{n=1}^{\infty} n!x^n$$

呢？我们有

$$\lim_{n\to\infty}\left|\frac{(n+1)!x^{n+1}}{n!x^n}\right| = \lim_{n\to\infty}\left|\frac{(n+1)!}{n!}\frac{x^{n+1}}{x^n}\right| = \lim_{n\to\infty}(n+1)|x|.$$

最后的极限是什么? 若 $x=0$, 则当 $n\to\infty$ 时, $0(n+1)=0$ 的极限当然为 0. (你可能注意到了, 这种情况下的 x^{n+1}/x^n 并没定义!) 然而, 对其他的任何 x 值, 我们就有点晕了 —— 极限是 ∞, 肯定大于 1. 我们得出结论, 级数只在 $x=0$ 时收敛 (要知道, 级数必在 $x=a$ 处收敛, 在这个例子中 a 为 0). 所以收敛半径为 0, 且图像如图 26-9 所示.

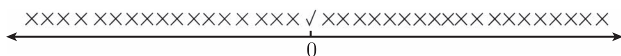


图 26-9

现在考虑

$$\sum_{n=1}^{\infty}\frac{(-2)^n}{\sqrt{n}}(x-7)^n.$$

这是一个 $a=7$ 的幂级数, 所以该点肯定在收敛区域的中心. 不管怎样, 通过讨论, 我们有

$$\lim_{n\to\infty}\left|\frac{\dfrac{(-2)^{n+1}(x-7)^{n+1}}{\sqrt{n+1}}}{\dfrac{(-2)^n(x-7)^n}{\sqrt{n}}}\right| = \lim_{n\to\infty}\left|\frac{(-2)^{n+1}}{(-2)^n}\frac{(x-7)^{n+1}}{(x-7)^n}\sqrt{\frac{n}{n+1}}\right|$$

$$= 2|x-7|.$$

所以幂级数在 $2|x-7|<1$ 时绝对收敛, 在 $2|x-7|>1$ 时发散. 两边除以 2, 可知级数在 $|x-7|<\frac{1}{2}$ 时收敛, 在 $|x-7|>\frac{1}{2}$ 时发散. 故收敛半径为 $\frac{1}{2}$, 图像如图 26-10 所示.

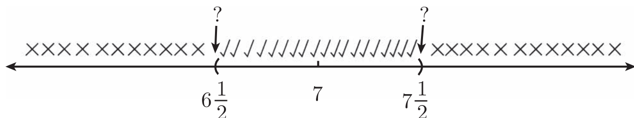


图 26-10

我们仍需讨论端点. 试一下 $x=7\frac{1}{2}$, 则级数为

$$\sum_{n=1}^{\infty}\frac{(-2)^n}{\sqrt{n}}\left(7\frac{1}{2}-7\right)^n = \sum_{n=1}^{\infty}\frac{(-2)^n}{\sqrt{n}}\frac{1}{2^n} = \sum_{n=1}^{\infty}\frac{(-1)^n}{\sqrt{n}}.$$

要确保, 你意识到了为什么 $(-2)^n/2^n$ 能化简到 $(-1)^n$. 不管怎样, 我把证明最后这个级数条件收敛 (用交错级数判别法) 而非绝对收敛 (用 p 判别法) 留给你自行完成. 现在, 当 $x=6\frac{1}{2}$, 可得

$$\sum_{n=1}^{\infty}\frac{(-2)^n}{\sqrt{n}}\left(6\frac{1}{2}-7\right)^n = \sum_{n=1}^{\infty}\frac{(-2)^n}{\sqrt{n}}\left(-\frac{1}{2}\right)^n = \sum_{n=1}^{\infty}\frac{(-2)^n}{\sqrt{n}}\frac{1}{(-2)^n} = \sum_{n=1}^{\infty}\frac{1}{\sqrt{n}},$$

发散. 我们得出结论, 幂级数在 $6\frac{1}{2} < x < 7\frac{1}{2}$ 时绝对收敛, 在 $x = 7\frac{1}{2}$ 时条件收敛, 其他情况发散, 完整图示见图 26-11.

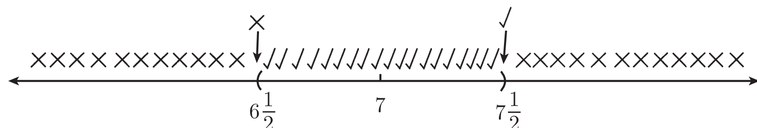


图 26-11

考虑级数

$$\sum_{n=1}^{\infty} \frac{3^n}{2^{n^2}}(x+2)^n.$$

该级数因为复杂的因子 2^{n^2} 使其更适合运用根式判别法. 你可以用比式判别法求出结果, 但根式判别法更好. 考虑第 n 项绝对值的 n 次方根的极限：

$$\lim_{n\to\infty}\left|\frac{3^n}{2^{n^2}}(x+2)^n\right|^{1/n} = \lim_{n\to\infty}\frac{(3^n)^{1/n}}{(2^{n^2})^{1/n}}(|x+2|^n)^{1/n} = \lim_{n\to\infty}\frac{3}{2^n}|x+2|.$$

现在无论 x 取何值, 极限都等于 0, 小于 1; 根据根式判别法, 幂级数对所有 x 都绝对收敛, 即收敛半径是 ∞, 图像如图 26-12 所示.

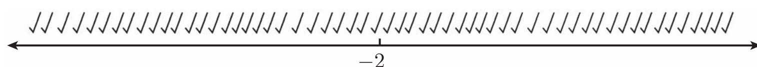


图 26-12

在进入下一节前, 这里还有最后一点说明：当收敛半径为正时, 可能在两端点都收敛, 或在两端点都不收敛, 或只在左端点收敛, 或只在右端点收敛. 我们在前面见过了所有这四种可能.

26.2 合成新的泰勒级数

我们来看一些求泰勒级数的方法. 求给定函数 f 关于 $x = a$ 的泰勒级数的一个方法, 是像 25.2 节那样直接用公式. 为了运用公式需要求 f 的所有导数, 至少是在 $x = a$ 的所有导数. 对大多数函数来说, 这是一件令人厌烦的事. 通常, 一个较好的办法是用一些常见的泰勒级数来合成新的泰勒级数. 当然, 首先你需要知道一些泰勒级数! 下面 5 个麦克劳林级数 (关于 $x = 0$ 的泰勒级数) 是非常有用的.

(1) 对应 $f(x) = \mathrm{e}^x$：

$$\boxed{\mathrm{e}^x = \sum_{n=0}^{\infty}\frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots}$$

对所有实数 x 都成立.

(2) 对应 $f(x) = \sin(x)$：

$$\boxed{\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots}$$

对所有实数 x 都成立.

(3) 对应 $f(x) = \cos(x)$：

$$\boxed{\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots}$$

对所有实数 x 都成立.

(4) 对应 $f(x) = 1/(1-x)$：

$$\boxed{\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots}$$

只对 $-1 < x < 1$ 成立.

(5) 对应 $f(x) = \ln(1+x)$ 或 $f(x) = \ln(1-x)$：

$$\boxed{\begin{aligned} \ln(1+x) &= \sum_{n=1}^{\infty} -\frac{(-1)^n x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \\ \ln(1-x) &= \sum_{n=1}^{\infty} -\frac{x^n}{n} = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots \end{aligned}}$$

对 $-1 < x < 1$ 成立. (其实, 第一个公式也对 $x = 1$ 成立, 第二个公式也对 $x = -1$ 成立, 不过这个有点复杂了!)

至今为止, 我们已经证明了公式 (1) 和 (3) (见 24.2.3 节) 和 (4) (见 22.2 节). 后面的 26.2.2 节和 26.2.3 节将分别讨论 (2) 和 (5).

无论如何, 我假设你已经学过了这 5 个级数. 下面介绍如何通过对它们进行操作来得到新的幂级数.①

26.2.1 代换和泰勒级数

最有用的方法就是做代换. 在麦克劳林级数中, 你可以将 x 换为 x^n 的倍数来得到一个新的麦克劳林级数, 其中 n 是一个整数. 例如, 我们知道



$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

对任意 x 都成立, 故若想求 $f(x) = e^{x^2}$ 的麦克劳林级数, 只需将上述级数的 x 换成 x^2, 可得

$$e^{x^2} = 1 + x^2 + \frac{(x^2)^2}{2!} + \frac{(x^2)^3}{3!} + \frac{(x^2)^4}{4!} + \cdots,$$

① 这些方法的证明不在本书讨论范围内.

并可化简为

$$e^{x^2} = 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \frac{x^8}{4!} + \cdots .$$

由于原级数对任意 x 都成立, 这个级数也一样.

我们来看另一个常见的例子：$f(x) = 1/(1+x^2)$ 的麦克劳林级数是什么？要求解该题, 我们从下面的几何级数开始：

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots ,$$

它对 $-1 < x < 1$ 成立, 然后将 x 换成 $-x^2$ 可得

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-x^2)^n = \sum_{n=0}^{\infty}(-1)^n x^{2n} = 1 - x^2 + x^4 - x^6 + \cdots ,$$

它对 $-1 < -x^2 < 1$ 成立. 注意, 我们将这个“成立”的不等式中的 x 换成了 $-x^2$. 在此, 这并不重要, 因为不等式最后可化简为 $-1 < x < 1$. 但是假设我们要求 $1/(1+2x^2)$ 的麦克劳林级数, 则需将 x 换成 $-2x^2$, 此时可得

$$\frac{1}{1+2x^2} = \sum_{n=0}^{\infty}(-2x^2)^n = \sum_{n=0}^{\infty}(-1)^n 2^n x^{2n} = 1 - 2x^2 + 4x^4 - 8x^6 + \cdots ,$$

但它只对 $-1 < -2x^2 < 1$ 成立. 可以确信, 该不等式可化为 $-1/\sqrt{2} < x < 1/\sqrt{2}$. (这里的所有级数都是几何级数.)

假设现在从下面的等式开始, 该等式对所有的 x 都成立：

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots .$$

右边是 $\sin(x)$ 的麦克劳林级数, 或可看作 $\sin(x)$ 关于 $x = 0$ 的泰勒级数. 若将 x 用 $(x-18)$ 代换, 则得到一个关于 $x = 18$ 的泰勒级数：

$$\sin(x-18) = (x-18) - \frac{(x-18)^3}{3!} + \frac{(x-18)^5}{5!} - \frac{(x-18)^7}{7!} + \cdots .$$

右边不是 $\sin(x)$ 关于 $x = 18$ 的泰勒级数, 因为左边不再是 $\sin(x)$, 而是 $\sin(x-18)$. 所以我们的代换也改变了原函数. 我们其实求出了 $\sin(x-18)$ 关于 $x = 18$ 的泰勒级数. 为了求出 $\sin(x)$ 关于 $x = 18$ 的泰勒级数, 需要用到泰勒定理中的公式. (我们在 25.2 节末见过类似的问题.)

上面这个例子告诉我们, 若将 x 换为 $(x-a)$, 则得到关于 $x = a$ 的泰勒级数而不是麦克劳林级数, 且函数也变了. 这还是有用的. 例如, 为了求 $\ln(x)$ 关于 $x = 1$ 的泰勒级数, 我们从前一节的公式

$$\ln(1+x) = \sum_{n=1}^{\infty} -\frac{(-1)^n x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots , \quad -1 < x < 1.$$

开始. 现在, 将 x 换为 $(x-1)$, 则 $\ln(1+x)$ 变为 $\ln(1+(x-1))$, 即 $\ln(x)$, 则我们得到

$$\ln(x)=\sum_{n=1}^{\infty}-\frac{(-1)^n(x-1)^n}{n}=(x-1)-\frac{(x-1)^2}{2}+\frac{(x-1)^3}{3}-\frac{(x-1)^4}{4}+\cdots,$$
$$-1<(x-1)<1.$$

注意, 我同样将原不等式 $-1<x<1$ 中的 x 换为 $(x-1)$, 得到 $-1<(x-1)<1$. 这个不等式看起来有些蠢, 故各项加 1, 得到 $0<x<2$. 最后可得

$$\ln(x)=\sum_{n=1}^{\infty}-\frac{(-1)^n(x-1)^n}{n}=(x-1)-\frac{(x-1)^2}{2}+\frac{(x-1)^3}{3}-\frac{(x-1)^4}{4}+\cdots,$$
$$0<x<2.$$

这里用了 $\ln(1+x)$ 的麦克劳林级数得到 $\ln(x)$ 关于 $x=1$ 的泰勒级数.

代换方法也可以用于求泰勒多项式, 不过要注意写对阶数. 例如, 若取 $f(x)=\mathrm{e}^x$ 和 $a=0$, 则 3 阶泰勒多项式为

$$P_3(x)=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}.$$

若 $g(x)=\mathrm{e}^{x^2}$, 将上述多项式的 x 换为 x^2, 则认为 g 的 3 阶泰勒多项式为

$$P_3(x)=1+x^2+\frac{x^4}{2!}+\frac{x^6}{3!}.$$

这是错的. 它其实是 g 关于 $x=0$ 的 6 阶泰勒多项式, 所以左边应为 $P_6(x)$ 而不是 $P_3(x)$. 为了得到 $P_3(x)$ 的正确公式, 只需去掉所有次数大于 3 的项, 即为 $P_3(x)=1+x^2$. 当然, 它也是 $P_2(x)$! 当心, 不要看作次数哦! 那可是阶数. (至少, 你想通过微积分这门课并取得学位 …… 好吧, 我发誓再也不使用双关语[1]了.)

26.2.2 泰勒级数求导

若一个幂级数收敛于开区间 (a,b) 上可导的函数 f, 则可以通过对幂级数逐项求导, 得到一个在相同区间上收敛于 $f'(x)$ 的新幂级数. 在端点 a 和 b 的情况比较棘手：求导后的级数可能发散, 即使原级数是收敛的[2]所以要单独讨论端点.

我们的第一个例子是求 $\sin(x)$ 的麦克劳林级数, 假设已知 $\cos(x)$ 的麦克劳林级数为



$$\cos(x)=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\frac{x^8}{8!}-\cdots,$$

该公式对所有 x 都成立. (这个我们已在 24.2.3 节证明.) 若两边同时求导, 右边逐项求导, 可得

$$-\sin(x)=-\frac{2x}{2!}+\frac{4x^3}{4!}-\frac{6x^5}{6!}+\frac{8x^7}{8!}-\cdots.$$

① 这里指上一句的阶数 order 一词, 该词也有 “命令” 之意. —— 编者注

② 若求导后的级数在一个 (或两个) 端点处收敛, 则原级数也在那里收敛.

为了处理左边的负号, 两边同乘 -1. 不过还需要做另一步化简. 我们要处理形如 $2/2!$、$4/4!$、$6/6!$ 和 $8/8!$ 的量. 先来考虑 $4/4!$, 由于 $4!$ 实为 $3!\times 4$, 所以可通过消掉因子 4 而将 $4/4!$ 化简为 $1/3!$. 类似地, $6! = 5!\times 6$, 故有 $6/6! = 1/5!$, 同样 $8! = 7!\times 8$, 所以 $8/8! = 1/7!$. 综上, 上面的等式变为

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots.$$

由于 $\cos(x)$ 的级数对所有 x 都成立, 所以上述求导后的级数也如此. 即, $\sin(x)$ 的麦克劳林级数由上式给出, 且对所有 x 都成立. 这就证明了 26.2 节的公式 (2).

这里是幂级数求导的另一个例子. 假定欲求 $f(x) = 1/(1+x)^2$ 的麦克劳林级数. 最好的方法是从 $1/(1+x)$ 的级数开始, 该级数是通过将标准几何级数 (前面的公式 (4)) 的 x 换为 $-x$ 而得到的:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \cdots;$$

对 $-1 < x < 1$ 成立. 然后两边求导, 右边逐项求导, 可得

$$-\frac{1}{(1+x)^2} = 0 - 1 + 2x - 3x^2 + 4x^3 - \cdots.$$

剩下的就是两边同时取负, 得

$$\frac{1}{(1+x)^2} = 1 - 2x + 3x^2 - 4x^3 + \cdots = \sum_{n=0}^{\infty}(-1)^n(n+1)x^n,$$

对 $-1 < x < 1$ 成立. (你需要验证, 带求和号的表达式是正确的, 且级数在端点 $x = \pm 1$ 处不收敛.)

同样, 你可以将这些方法用于泰勒多项式, 还是要注意阶数. 由于多项式求导使得次数减 1, 所以求导后的泰勒多项式的阶比原多项式的阶小 1. 例如, $1/(1+x)$ 关于 0 的 3 阶泰勒多项式是 $1 - x + x^2 - x^3$, 如前一个例子. 若求导并乘以 -1, 则 $1/(1+x)^2$ 关于 0 的二阶泰勒多项式为 $1 - 2x + 3x^2$.

26.2.3　泰勒级数求积分

我们还可以对泰勒级数逐项求积分. 新的级数与原级数收敛区间一样 (收敛区间的端点除外). 若用的是不定积分, 别忘了常数! 我们来看一些例子. 首先, 证明 $\ln(1-x)$ 的公式, 这是 26.2 节的公式 (5), 不过没证过:

$$\ln(1-x) = \sum_{n=1}^{\infty} -\frac{x^n}{n} = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots, \quad -1 < x < 1.$$

我们将用到几何级数的公式, 即 26.2 节的公式 (4):

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots, \quad -1 < x < 1.$$

然后对每一项关于 x 求积分：

$$\int \frac{1}{1-x}\mathrm{d}x = \int \sum_{n=0}^{\infty} x^n \mathrm{d}x = \int (1+x+x^2+x^3+\cdots)\mathrm{d}x.$$

(注意, 我既用了求和号, 也用了展开式, 不过你一般只用其中之一.) 现在逐项求积分：

$$-\ln(1-x) = C + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1} = C + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots.$$

这里最好将常数放在前面, 而不以 $+C$ 的方式放在后面, 因为常数是幂级数的零次项. 现在我们要求出 C 的值. 最好的方法是代入 $x=0$, 由此可得

$$-\ln(1-0) = C + 0 + \frac{0^2}{2} + \frac{0^3}{3} + \frac{0^4}{4} + \cdots,$$

化简后得 $C=0$. 将其代入并两边取负, 则得到前面的 $\ln(1-x)$ 的级数：

$$\ln(1-x) = \sum_{n=1}^{\infty} -\frac{x^n}{n} = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots.$$

由于原级数 ($1/(1+x)$ 的级数) 对 $-1<x<1$ 收敛, 故积分后的级数 (即 $-\ln(1-x)$ 的级数, 进而对 $\ln(1-x)$ 的级数) 也对 $-1<x<1$ 收敛. 其实, $\ln(1-x)$ 的级数当 $x=-1$ 时也收敛, 不过如我所说, 逐项积分以后的幂级数并未给出收敛区间端点的任何信息. 现在, 可将 26.2 节公式 (5) 中的 x 代换为 $-x$, 得到 $\ln(1+x)$ 的展开式.

另一个例子：如何求 $\tan^{-1}(x)$ 的麦克劳林级数? 不断求导是很痛苦的 (试试看就知道了!) 但我们可以更灵活一点, 对已知的级数求积分. 我们来看一下, $\tan^{-1}(x)$ 是 $1/(1+x^2)$ 的一个反导数, 我们在 26.2.1 节得知



$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots, \quad -1<x<1.$$

现在可以两边求积分, 得到

$$\int \frac{1}{1+x^2}\mathrm{d}x = \int (1 - x^2 + x^4 - x^6 + \cdots)\mathrm{d}x.$$

右边逐项求积可得

$$\tan^{-1}(x) = C + x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots.$$

代入 $x=0$ 来求 C：

$$\tan^{-1}(0) = C + 0 - \frac{0^3}{3} + \frac{0^5}{5} - \frac{0^7}{7} + \cdots,$$

化简为 $C=\tan^{-1}(0)=0$. 故, 我们有

$$\tan^{-1}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}.$$

(确信右边的求和号形式是正确的.) 由于 $1/(1+x^2)$ 的原级数在 $-1<x<1$ 时收敛, 所以 $\tan^{-1}(x)$ 的级数也在 $-1<x<1$ 时收敛.[①]

我们来看一个定积分的例子. 假定函数 f 定义为

$$f(x)=\int_0^x \sin(t^3)\mathrm{d}t.$$

它的麦克劳林级数是什么? 我们应该从求 $\sin(t^3)$ 的级数开始. 为此, 对 $\sin(x)$ 的麦克劳林级数做换元 $x=t^3$, 可得

$$\begin{aligned}\sin(t^3)&=t^3-\frac{(t^3)^3}{3!}+\frac{(t^3)^5}{5!}-\frac{(t^3)^7}{7!}+\cdots\\&=t^3-\frac{t^9}{3!}+\frac{t^{15}}{5!}-\frac{t^{21}}{7!}+\cdots.\end{aligned}$$

由于 $\sin(x)$ 的级数对所有实数 x 均成立, 则 $\sin(t^3)$ 的级数对所有实数 t 都成立. 现在两边可同时求 0 到 x 的积分, 得

$$f(x)=\int_0^x \sin(t^3)\mathrm{d}t=\int_0^x\left(t^3-\frac{t^9}{3!}+\frac{t^{15}}{5!}-\frac{t^{21}}{7!}+\cdots\right)\mathrm{d}t.$$

对右边逐项求积分, 可得

$$\begin{aligned}f(x)&=\left(\frac{t^4}{4}-\frac{t^{10}}{10\cdot 3!}+\frac{t^{16}}{16\cdot 5!}-\frac{t^{22}}{22\cdot 7!}+\cdots\right)\Bigg|_0^x\\&=\frac{x^4}{4}-\frac{x^{10}}{10\cdot 3!}+\frac{x^{16}}{16\cdot 5!}-\frac{x^{22}}{22\cdot 7!}+\cdots,\end{aligned}$$

对所有实数 x 都成立. (你应该试着将这个级数写成求和号的形式, 答案在 26.3 节给出.)

也可将上述积分方法用于泰勒多项式, 这次泰勒多项式的阶要加 1.

26.2.4　泰勒级数相加和相减

若已知两个函数 f 和 g 关于 $x=a$ 的泰勒级数, 则和式 $f(x)+g(x)$ 的泰勒级数显然是两个泰勒级数之和, 这至少对于两泰勒级数收敛区间的交集是成立的. 差 $f(x)-g(x)$ 遵循相同规则. 在实践中唯一需要做的事就是合并同类项, 然后关注所得级数在哪里收敛. 例如, $\sin(x)-\mathrm{e}^x$ 的麦克劳林级数为

$$\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots\right)-\left(1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\frac{x^6}{6!}+\frac{x^7}{7!}+\cdots\right),$$

这里需要化简. 消减后, 至少到 7 阶的级数为

① 其实, 根据交错级数判别法, $\tan^{-1}(x)$ 的级数在 $x=1$(或 $x=-1$) 时也收敛, 最后可得一个漂亮的公式

$$1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots=\tan^{-1}(1)=\frac{\pi}{4}.$$

$$-1-\frac{x^2}{2!}-\frac{2x^3}{3!}-\frac{x^4}{4!}-\frac{x^6}{6!}-\frac{2x^7}{7!}-\cdots,$$

由于 $\sin(x)$ 和 e^x 的级数对所有 x 成立, 所以 $\sin(x)-e^x$ 的级数也一样.

若讨论泰勒多项式, 则需要注意阶数取两个阶数中的较小者. 例如, 我们知道 $1/(1-x)$ 关于 0 的三阶泰勒多项式为

$$1+x+x^2+x^3,$$

而 e^x 关于 0 的四阶泰勒多项式为

$$1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}.$$

若 $f(x)=1/(1-x)+e^x$, 求它关于 0 的泰勒多项式, 则取上述两个多项式的和是不对的. 问题出在 e^x 的多项式有四阶项, 但 $1/(1-x)$ 没有四阶项. 这好比是拿苹果和橘子这样两种无法相比的东西做比较. 你应该将 e^x 多项式的四阶项略去来得到三阶泰勒多项式

$$1+x+\frac{x^2}{2!}+\frac{x^3}{3!}.$$

现在可将 $1+x+x^2+x^3$ 加到上面的多项式, 得到 $1/(1-x)+e^x$ 关于 $x=0$ 的三阶泰勒多项式

$$(1+x+x^2+x^3)+\left(1+x+\frac{x^2}{2!}+\frac{x^3}{3!}\right),$$

化简可得

$$2+2x+\frac{3x^2}{2}+\frac{7x^3}{6}.$$

26.2.5 泰勒级数相乘

你也可以将两个泰勒级数相乘, 从而得到一个收敛于两个函数之积的新级数, 至少该级数在两个泰勒级数收敛区域的交集收敛. 用求和号形式书写这些会很乱, 且通常会有两个求和号. 一般地, 大家只关注级数的前面几项. 例如, 求 $f(x)=e^x\sin(x)$ 的三阶及以下的麦克劳林级数. 欲求该问题, 写出 e^x 和 $\sin(x)$ 的三阶及以下的级数, 相乘, 然后略去所有大于三阶的项:



$$\begin{aligned}e^x\sin(x)&=\left(1+x+\frac{x^2}{2}+\frac{x^3}{6}+\cdots\right)\left(x-\frac{x^3}{6}+\cdots\right)\\&=1\left(x-\frac{x^3}{6}\right)+x(x)+\frac{x^2}{2}(x)+\cdots\\&=x+x^2+\frac{x^3}{3}+\cdots\end{aligned}$$

有一个略去无用项的技巧. 例如, 分别略去第一个和与第二个和中的项 x 和 $-x^3/6$ 的乘积, 因为我意识到它们之积会得到一个含 x^4 的项, 而这不是我关心的项, 因为我只需要到三阶的项. 若我要关心到四阶的项, 则必然要关注更多的项.

事实上, 不要把注意力集中在次数大于原函数级数的阶的项, 这点很重要. 例如, 取 e^x 关于 0 的二阶泰勒多项式

$$1+x+\frac{x^2}{2};$$

现在令其与 e^{-x} 关于 0 的二阶泰勒多项式

$$1-x+\frac{x^2}{2}$$

相乘, 得到

$$\left(1+x+\frac{x^2}{2}\right)\left(1-x+\frac{x^2}{2}\right),$$

化简为

$$1+\frac{x^4}{4}.$$

若你说, 它是乘积 $(e^x)(e^{-x})$ 关于 0 的四阶泰勒多项式, 那就大错特错了! 毕竟, 两函数之积是 1, 故它的所有泰勒多项式都是 1. 正确的做法是, 略去积中所有次数大于 2 的项. 毕竟, 我们只是从二阶多项式开始, 怎么能期望将这两个多项式相乘就得到更高阶呢? 在上面的多项式 $1+x^4/4$ 中, 项 $x^4/4$ 的次数大于 2, 故不准确, 应该略去. 多项式的二阶多项式为 1, 这就是你能从两个二阶泰勒多项式的积中得到的所有结论. 不要将精力都集中在更高次数上, 以免贪多嚼不烂而使自己骑虎难下.

26.2.6　泰勒级数相除



你可以用长除法来做与除法一样的事. 方法是略掉不关心的项. 例如, 为了求 $f(x)=\sec(x)$ 的四阶麦克劳林级数, 首先将 $\sec(x)$ 写为 $1/\cos(x)$, 然后与多项式一样做长除法. 这里的主要区别是, 你应该将各项按次数递增的顺序写, 而不是平常的递减顺序写. 由于我们关心四阶及以下的项, 所以将

$$\cos(x)=1-\frac{x^2}{2}+\frac{x^4}{24}-\cdots$$

用在 $1/\cos(x)$ 的长除法中:

$$
\begin{array}{r|l}
 & 1 \qquad + \tfrac{1}{2}x^2 \qquad + \tfrac{5}{24}x^4 + \cdots \\
\hline
1+0x-\tfrac{1}{2}x^2+0x^3+\tfrac{1}{24}x^4-\cdots & 1+0x+0x^2+0x^3+0x^4+\cdots \\
 & \underline{1+0x-\tfrac{1}{2}x^2+0x^3+\tfrac{1}{24}x^4+\cdots} \\
 & \tfrac{1}{2}x^2+0x^3-\tfrac{1}{24}x^4+\cdots \\
 & \underline{\tfrac{1}{2}x^2+0x^3-\tfrac{1}{4}x^4+\cdots} \\
 & \tfrac{5}{24}x^4+\cdots
\end{array}
$$

所以 $\sec(x)$ 的麦克劳林级数是 $1+x^2/2+5x^4/24+\cdots$, 直到四阶项.



若我们欲求 $\tan(x)$ 的四阶麦克劳林级数, 可类似求解, 因为 $\tan(x)=\sin(x)/\cos(x)$. 利用 $\sin(x)=x-x^3/6+\cdots$ 和 $\cos(x)=1-x^2/2+x^4/24-\cdots$, 除法如下:

$$1 + 0x - \frac{1}{2}x^2 + 0x^3 + \frac{1}{24}x^4 - \cdots \overline{\Big)\, 0 + x + 0x^2 - \frac{1}{6}x^3 + 0x^4 + \cdots}$$



计算请自行完成. 通过计算可知, 当含有四阶及以下项时, $\tan(x) = x + x^3/3 + \cdots$ (注意这里四阶项为 0).

以上论述表明你可能不需要连续求导, 而利用泰勒级数公式来求泰勒级数. 若幸运的话, 可以用 5 个基本级数, 外加一个或多个如换元、求导、积分、相加、相减、相乘和相除的方法来求级数.

26.3 利用幂级数和泰勒级数求导

回忆 $f(x)$ 关于 $x = a$ 的泰勒级数的第 n 项系数公式：

$$a_n = \frac{f^{(n)}(a)}{n!}.$$

两边同乘 $n!$ 得到：

$$\boxed{f^{(n)}(a) = n! \times a_n.}$$

用语言描述, 意思是

$$\boxed{f^{(n)}(a) = n! \times \ (f(x) \text{ 在 } x = a \text{ 处的泰勒级数中 } (x-a)^n \text{ 的系数})}$$

所以若知道一个函数关于某点 a 的泰勒级数, 就可以很容易地求得该函数在 a 点的导数. 这就是你的全部所得! 这里并没有任何关于其他 x 值的导数值的信息, 只有 $x = a$. (其实, 为求第 n 阶导数, 只需要一个在 $x = a$ 的 n 阶或更高阶的泰勒多项式, 而不是整个泰勒级数.)



为了应用上面的方程, 需先求给定函数的一个合适的泰勒级数. 前几节的方法也很有用的. 例如, 假设 $f(x) = \mathrm{e}^{x^2}$, 我们欲求 $f^{(100)}(0)$ 和 $f^{(101)}(0)$. 我们从求 e^{x^2} 的麦克劳林级数开始：

$$\mathrm{e}^{x^2} = \sum_{n=0}^{\infty} \frac{(x^2)^n}{n!} = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!} = 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \cdots.$$

根据前面方框中的公式,

$$f^{(100)}(0) = 100! \times (\text{上述麦克劳林级数中 } x^{100} \text{ 的系数}).$$

那么麦克劳林级数中 x^{100} 的系数是什么？看上面的麦克劳林级数, 可知系数就是 $1/(50!)$, 或者更正式地说, 你能够算出 n 的什么值对应 x^{100}. 特别地, 我们想确定 x^{100} 的倍数 $x^{2n}/n!$, 而这就意味着 $2n = 100$, 所以 $n = 50$, 对应的项为 $x^{100}/(50!)$. 故系数是 $1/(50!)$. 于是

$$f^{(100)}(0) = 100! \times \frac{1}{50!} = \frac{100!}{50!}.$$

(不要犯将最后的表达式化简为 2! 的错误, 阶乘不是这样算的.) 现在想一想, $f^{(101)}(0)$ 又怎么求呢? 它等于上述级数中 x^{101} 系数的 101! 倍. 那个系数是什么? 等一下, 级数中没有奇次幂! 换一种方式思考, 什么样的 n 值对应 x^{101}? 需要解 $2n = 101$, 但 n 必须为整数, 所以没有幂 x^{101}. 那就意味着 x^{101} 的系数为 0, 所以

$$f^{(101)}(0) = 101! \times 0 = 0.$$

好吧, 我们来看一个更难的例子. 在 26.2.3 节, 我们发现函数

$$f(x) = \int_0^x \sin(t^3)\mathrm{d}t$$

的麦克劳林级数是

$$\frac{x^4}{4} - \frac{x^{10}}{10 \cdot 3!} + \frac{x^{16}}{16 \cdot 5!} - \frac{x^{22}}{22 \cdot 7!} + \cdots,$$

这个级数对所有的 x 都收敛于 $f(x)$. 现在要问: $f^{(50)}(0)$ 是什么? $f^{(52)}(0)$ 呢? 为了求出这些, 我们需要知道前面 $f(x)$ 的级数中 x^{50} 和 x^{52} 的系数. 要知道, $f^{(50)}(0)$ 是 $f(x)$ 麦克劳林级数中 x^{50} 系数的 50! 倍, 当然, 除了处处用 52 代替 50 之外, $f^{(52)}(0)$ 也同理.

为了求上面级数中 x^{50} 和 x^{52} 的系数, 需要将级数写出至足够长以便于理解. 更好的方法是将级数用求和号表示. 之前我已经让你练习过, 这里是相应的做法. 注意 x 的幂为 4, 10, 16, 22, $\cdots$. 这意味着幂次从 4 开始, 每次增长 6. 所以, 指数为 $6n+4$, 其中 n 取值为 0, 1, 2, 3, $\cdots$. 现在来看分母, 它是 $6n+4$ 与某奇数阶乘的乘积. 其中奇数为 1, 3, 5, 7, $\cdots$, 故分母是 $(6n+4)(2n+1)!$. 最后, 各项以正项开始, 正负交错, 所以应该还有 $(-1)^n$. 现在, 我们得到

$$f(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{6n+4}}{(6n+4)(2n+1)!}.$$

现在来求 x^{50} 和 x^{52} 的系数. 对前者, 解 $6n+4=50$ 得到 $n = 23/3$, 它不是整数, 所以 x^{50} 的系数为 0. 意味着

$$f^{(50)}(0) = 50! \times (x^{50} \text{ 的系数}) = 50! \times 0 = 0.$$

另外对 x^{52}, 解 $6n+4=52$ 得到 $n=8$, 故我们可以通过观察 $n=8$ 时的结果来确定 x^{52} 的系数. 和式中 $n=8$ 的项是

$$\frac{(-1)^8 x^{6\times 8+4}}{(6\times 8+4)(2\times 8+1)!} = \frac{x^{52}}{52\times 17!},$$

所以系数是 $1/(52\times 17!)$. 最后,

$$f^{(52)}(0) = 52! \times (x^{52}\text{的系数}) = 52! \times \frac{1}{52\times 17!} = \frac{51!}{17!}.$$

注意, 这里做了一个小的相消：$52!/52 = 51!$, 在继续之前要明白这么做的正确性!

有时, 一个函数已经被关于 $x=a$ 的一个幂级数定义, 你可能需要求这个函数在 a 处的某些导数. 这个甚至比前面的例子容易些, 因为不用先求泰勒级数. 例如, 假设 $f(x)$ 定义为



$$f(x)=\sum_{n=0}^{\infty}\frac{(-1)^{n+1}n^3(x-6)^{3n}}{n!},$$

它对所有 x 都收敛 (为什么?). 假设要求 $f^{(300)}(6)$ 的值. 幂级数是关于 $x=6$ 的, 故利用公式

$$f^{(300)}(6)=300!\times(\text{上面级数中 }(x-6)^{300}\text{ 的系数}).$$

要知道系数的值, 应该求出 n 的什么值给出了正确的项. 看上面的级数, $(x-6)$ 的指数为 $3n$, 所以我们需要 $3n=300$ 的项. 因此 $n=100$, 代入后可以看到正确的项是

$$\frac{(-1)^{100+1}100^3(x-6)^{300}}{100!}=\frac{-1\,000\,000}{100!}(x-6)^{300}.$$

所以系数是 $-1\,000\,000/100!$. 要想使它更别致一点, 可以将 $100!$ 写成 $100\times 99!$ 并消掉 100, 得到系数为 $-10\,000/99!$. 总之, 这个给出了

$$f^{(300)}(6)=300!\times\frac{-10\,000}{99!}=-\frac{300!\times 10\,000}{99!}.$$



那要求 $f^{(301)}(6)$ 怎么办呢? 可以证明幂级数中没有 $(x-6)^{301}$ 这项, 所以答案为 0, 这部分留给你自己完成.

26.4 利用麦克劳林级数求极限

你也可以利用泰勒级数来求特定的极限. 特别地, 若你有极限

$$\lim_{x\to 0}\frac{f(x)}{g(x)},$$



其中当 $x=0$ 时, 分子分母都为 0, 则可以用洛必达法则; 然而, 若想求

$$\lim_{x\to 0}\frac{\mathrm{e}^{-x^2}+x^2\cos(x)-1}{1-\cos(2x^3)}$$

的值, 要是还那么做, 你会发疯的. 对分子分母求导一次可不好玩, 更不必说可能要求 6 次了 (结果确实如此). 所以, 正确的方法是用合适的麦克劳林级数中足够多的项来做替代. "足够多的项" 是什么意思? 我们希望能消去一些项, 且不想让分子或分母为 0. 我们先求到第 8 阶来试一下. 写出完整的麦克劳林级数, 首先, 因为

$$\mathrm{e}^x=1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}+\cdots,$$

用 $-x^2$ 代换 x, 得到

$$\mathrm{e}^{-x^2}=1-x^2+\frac{x^4}{2}-\frac{x^6}{6}+\frac{x^8}{24}-\cdots.$$

又因为

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{6!} + \cdots.$$

通过两边乘 x^2 可得 $x^2\cos(x)$ 的级数：

$$x^2\cos(x) = x^2 - \frac{x^4}{2} + \frac{x^6}{24} - \frac{x^8}{6!} + \cdots.$$

若我们回到 $\cos(x)$ 的级数并用 $2x^3$ 代换 x, 可得

$$\cos(2x^3) = 1 - \frac{(2x^3)^2}{2} + \frac{(2x^3)^4}{24} - \cdots = 1 - 2x^6 + \frac{2}{3}x^{12} - \cdots,$$

这里我们甚至不需要最后那项, 更不必理会任何次数更高的项, 因为我们只决定到 8 阶. 当然, 把它放在其中也不会有坏处, 所以我们留下了它. 总之, 若将所有这些联系在一起, 分子就是

$$\begin{aligned}
&\mathrm{e}^{-x^2} + x^2\cos(x) - 1 \\
&= \left(1 - x^2 + \frac{x^4}{2} - \frac{x^6}{6} + \frac{x^8}{24} - \cdots\right) + \left(x^2 - \frac{x^4}{2} + \frac{x^6}{24} - \frac{x^8}{6!} + \cdots\right) - 1 \\
&= -\frac{1}{8}x^6 - \left(\frac{1}{24} - \frac{1}{720}\right)x^8 + \cdots,
\end{aligned}$$

而分母变为

$$1 - \cos(2x^3) = 1 - \left(1 - 2x^6 + \frac{2}{3}x^{12} - \cdots\right) = 2x^6 - \frac{2}{3}x^{12} + \cdots.$$

现在代入极限, 我们有

$$\lim_{x\to 0} \frac{\mathrm{e}^{-x^2} + x^2\cos(x) - 1}{1 - \cos(2x^3)} = \lim_{x\to 0} \frac{-\frac{1}{8}x^6 - \left(\frac{1}{24} - \frac{1}{720}\right)x^8 + \cdots}{2x^6 + \frac{2}{3}x^{12} + \cdots}.$$

上下同时除以最低次项 x^6, 并代入 $x = 0$ 可知该极限等于

$$\lim_{x\to 0} \frac{-\frac{1}{8} + \left(\frac{1}{24} - \frac{1}{720}\right)x^2 + \cdots}{2 - \frac{2}{3}x^6 + \cdots} = \frac{-1/8}{2} = -\frac{1}{16}.$$

所以可知, 阶数大于 6 的项都不用写出来 (这就是为什么我从不烦心要化简 $1/24 - 1/720$ 的原因). 基本上, 若所有的项都消去了, 就意味着你没用到足够多的项; 如果还有一些项, 说明你已经写出了足够多的项并能继续下去. 如果最高只能写到 5 阶 (或更少), 又得到了 0/0, 那么就不会继续下去了.



我们再看一个例子：求

$$\lim_{x\to 0}\left(\frac{1}{\sin(x)} - \frac{1}{\mathrm{e}^x - 1}\right).$$

这看起来不像是个分式, 所以第一步要做些代数运算. 取公分母, 就像我们在 14.1.3 节中对洛必达类型 **B1** 的极限所做的一样, 将极限写成

$$\lim_{x\to 0}\frac{\mathrm{e}^x-1-\sin(x)}{\sin(x)(\mathrm{e}^x-1)}.$$

现在, 我们有

$$\mathrm{e}^x-1=x+\frac{x^2}{2}+\frac{x^3}{6}+\cdots,$$

和

$$\sin(x)=x-\frac{x^3}{6}+\cdots.$$

把这些代入, 极限变为

$$\lim_{x\to 0}\frac{\left(x+\frac{x^2}{2}+\frac{x^3}{6}+\cdots\right)-\left(x-\frac{x^3}{6}+\cdots\right)}{\left(x-\frac{x^3}{6}+\cdots\right)\left(x+\frac{x^2}{2}+\frac{x^3}{6}+\cdots\right)}$$

$$=\lim_{x\to 0}\frac{\frac{x^2}{2}+\frac{x^3}{3}+\cdots}{\left(x-\frac{x^3}{6}+\cdots\right)\left(x+\frac{x^2}{2}+\frac{x^3}{6}+\cdots\right)}.$$

当 $x\to 0$ 时, 还是最低次起决定作用. 为了说明这个, 上下同除以 x^2. 不过, 我们可以稍微变通一下：在分母上, 让两个因子都除以 x, 这与整个分母除以 x^2 一样. 极限变为

$$\lim_{x\to 0}\frac{\frac{1}{2}+\frac{x}{3}+\cdots}{\left(1-\frac{x^2}{6}+\cdots\right)\left(1+\frac{x}{2}+\frac{x^2}{6}+\cdots\right)}=\frac{1/2}{(1)(1)}=\frac{1}{2}.$$

同样, 写出其他的项不会有什么坏处 —— 这里我只用了三阶, 不过更高阶也行. 其实, 甚至三阶项也没有参与计算, 分母中只用到了一阶项. 除非你是心理学家或对这事有很好的直觉, 否则猜测需要多少项真是太难了. 所以, 用较多的项比用较少的项要好, 因为你总是可以稍后略去它们; 然而若用太少的项, 你甚至都解不出问题.

这是前面所有极限可行的真正原因：若 f 有最低次项为 a_Nx^N 的麦克劳林级数, 则



$$f(x)\sim a_Nx^N, \text{当 } x\to 0$$

我们在 21.4.5 节提过这个结论, 与极限比较判别法联系起来是有用的. 事实上, 上述等式甚至对 f 的麦克劳林级数关于 $x=0$ 不收敛的情况也是成立的. 所以没必要讨论完整的麦克劳林级数：最低阶且非 0 的 f 关于 $x=0$ 的泰勒多项式足以了. 只有一个条件, f 的第 $N+1$ 阶导数在 0 附近有界. 下面是完整过程：根据泰勒定

理, 我们有

$$f(x) = a_N x^N + R_N(x) = a_N x^N + \frac{f^{(N+1)}(c)}{(N+1)!}x^{N+1},$$

其中 c 介于 0 和 x 之间. 现在两边同时除以 $a_N x^N$ 可得

$$\frac{f(x)}{a_N x^N} = 1 + \frac{f^{(N+1)}(c)}{a_N(N+1)!}x.$$

右边量 $f^{(N+1)}(c)/(a_N(N+1)!)$ 的绝对值当 $x \to 0$ 时有界, 因为分母是常数且我们已经假设分子是有界的. 现在可用三明治定理来证明上述方程右边的最后一项在 $x \to 0$ 时趋于 0. 即,

$$\lim_{x\to 0} \frac{f(x)}{a_N x^N} = 1.$$

也就是说：

$$f(x) \sim a_N x^N, \text{ 当 } x \to 0$$



证毕. 怎样？我们不仅得到了一个利用极限比较判别法的便利工具, 而且证明了前面的极限都是成立的. 例如, 为了真正证明极限

$$\lim_{x\to 0} \frac{\mathrm{e}^x - 1 - \sin(x)}{\sin(x)(\mathrm{e}^x - 1)},$$

我们应该注意到 $\mathrm{e}^x - 1 - \sin(x)$ 有一个以 $x^2/2$ 开始的麦克劳林级数, 所以当 $x \to 0$ 时 $\mathrm{e}^x - 1 - \sin(x) \sim x^2/2$. 类似地, 当 $x \to 0$ 时 $\sin(x) \sim x$, 且当 $x \to 0$ 时 $\mathrm{e}^x - 1 \sim x$. 由于可以乘以或除以这些渐近关系 (但不能做加法或减法!), 因而可以说

$$\frac{\mathrm{e}^x - 1 - \sin(x)}{\sin(x)(\mathrm{e}^x - 1)} \sim \frac{x^2/2}{(x)(x)}, \text{ 当 } x \to 0.$$

右边就是 1/2, 所以我们证明了

$$\lim_{x\to 0} \frac{\mathrm{e}^x - 1 - \sin(x)}{\sin(x)(\mathrm{e}^x - 1)} = \frac{1}{2}.$$

在现实中, 上面的方法 (运用带 $+\cdots$ 符号的整个级数) 是被广泛接受的, 虽然严格上讲它只是围绕真正问题论述的. 真正发生的已在前面关于余项 R_N 的论证中给出.

第 27 章　参数方程和极坐标

迄今为止, 我们已经画过很多笛卡儿坐标系下形如 $y = f(x)$ 的方程的图像. 现在, 我们将从不同的角度来看问题：首先看一下当坐标 x 和 y 不直接相关而是通过一个公共参数相联系时会怎样; 接着看一下当将整个坐标系换成完全不同的形式时又会发生什么. 当然, 我们也会做一些计算. 下面是本章的计划：

- 参数方程、图和求切线;
- 极坐标与笛卡儿坐标的互换;
- 求极坐标曲线的切线;
- 求由极坐标曲线围成的面积.

27.1　参 数 方 程

当写下形如 $y = x^2\sin(x)$ 的方程时, 你是将 y 表示为了关于 x 的函数. 所以, 若已有 x 的特定值, 则通过将该 x 值代入方程可以很容易地求出相应的 y 值. 另一方面, 考虑关系 $x^2 + y^2 = 9$. 若已有一个特定的 x 值, 则你需要稍费点力气来求得相应的 y 值. 其实, 可能会有多个 y 值与给定的 x 值对应, 也可能一个都没有. 当然, 你可以写成 $y = \pm\sqrt{9 - x^2}$ 的形式, 意思是：如果 $-3 < x < 3$, 则有两个 y 值对应于 x; 但若 $x = \pm 3$, 则只有一个 y 值与之对应.

来看另一种方法：假设 x 和 y 都是另一个变量 t 的函数, 例如

$$x = 3\cos(t) \quad \text{和} \quad y = 3\sin(t).$$

这里是想让你将 x 看作关于 t 的函数; 若你愿意, 甚至可以写成 $x(t) = 3\cos(t)$ 的形式加以强调. 对 y 同理. 若选定 t 的值, 则可通过将该 t 值代入上面的方程求得相应的 x 和 y 值. 变量 t 被称为参数, 上述方程被称为*参数方程*.

上述这对参数方程的图像是什么样的呢？我们来试着描点. 与选择 x 值求得相应 y 值的一般方法不同, 我们选择一些 t 值, 并求得相应的 x 和 y 值. 为了描点, 只能采用 x 和 y 值, 因为没有 t 轴! 总之, 因为有三角函数, 所以我们应确保选定的值包含 π. 假定我们用了下面的 t 值.

t	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
x					
y					

我们用方程 $x = 3\cos(t)$ 和 $y = 3\sin(t)$ 算出了对应的 x 和 y 值, 便可填表如下.

t	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
x	3	$3\sqrt{3}/2$	$3/\sqrt{2}$	$3/2$	0
y	0	$3/2$	$3/\sqrt{2}$	$3\sqrt{3}/2$	3

例如, $t = 0$ 对应于点 $(3, 0)$, $t = \pi/6$ 对应于点 $(3\sqrt{3}/2, 3/2)$. 上面 5 个点如图 27-1 所示.

看似我们正在讨论中心在原点且半径为 3 的 1/4 圆. 这并不奇怪, 只要知道关于三角函数的知识! (当然, 对任意的 t 值, 有 $x^2 + y^2 = (3\cos(t))^2 + (3\sin(t))^2 = 9(\cos^2(t) + \sin^2(t)) = 9$.) 若继续上表, 直到 $t = \pi$, 就描述了半圆; 而若一直到 $t = 2\pi$, 则得到整个圆. 那再继续下去会发生什么呢? 你就会重新描述这个圆. 若从 $t = 0$ 开始向负方向继续, 那就是在沿着顺时针方向而不是逆时针方向走圆. 注意, 若在圆上选一点 (x, y), 并不是只有一个 t 值与该点对应, 而是有无穷多个 2π 倍于 t 值的数与之对应. 例如, 若 n 为任意整数, 则 $t = 2\pi n$ 对应于 $x = 3$ 和 $y = 0$, 即点 $(3, 0)$.

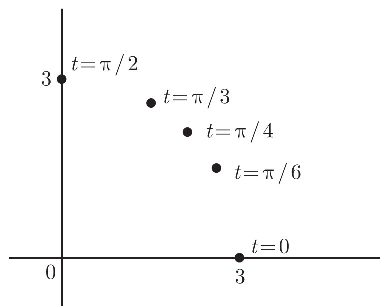


图　27-1

所以, 前面的这对参数方程描述了圆 $x^2 + y^2 = 9$, 至少 t 在一个足够大的区间 (例如, $[0, 2\pi)$) 里取值时是这样的. 你可以说

$$x = 3\cos(t) \quad 和 \quad y = 3\sin(t), 其中\ 0 \leqslant t \leqslant 2\pi$$

是 $x^2 + y^2 = 9$ 的参数化. 现在, 我问你: $x^2 + y^2 = 9$ 的图像与上面参数化的图像一样吗? 一样, 但也不一样. 当然, 两个图像看似是同一个圆, 不过参数化图像能告诉你更多信息: 圆是怎么画的. 若从 $t = 0$ 开始且连续移动到 $t = 2\pi$, 则你就可以从 $(3, 0)$ 开始并以不变的速度沿逆时针方向画, 直到回到起点.

通过观察, 整件事情就像蜗牛移动和离开时留下的粘液轨迹. 只是从轨迹并不足以看出蜗牛移动的方向 —— 它甚至可能往回走! 你也说不出它沿着轨迹移动时不同时间处的速度. (“蜗牛步伐” 不是它移动快慢的科学描述.) 借助参数化, 就像是知道了每一时刻蜗牛的位置一样, 能够知道方向和速度等其他信息.

那么, 上面的参数化是 $x^2 + y^2 = 9$ 的唯一可能的参数化吗? 当然不是. 还有很多其他方法可以画出相同的圆. 例如, 令 $x = 3\cos(2t)$ 和 $y = 3\sin(2t)$, 现在只需令 t 在 0 到 π 间取值就能包含整个圆, 并且这时的速度是原来的两倍. 亦或, 令 $x = 3\sin(t)$ 和 $y = 3\cos(t)$, 其中 $0 \leqslant t < 2\pi$. 现在又回到原来的速度了, 不过这次

是从 $(0,3)$ 开始以顺时针方向而不是逆时针方向运动. 可以通过描点来验证这些结论.

怎么求 $x^2+4y^2=9$ 的参数化? 画该方程的图像得到一个通过点 $(\pm 3,0)$ 和 $(0,\pm 3/2)$ 的椭圆. 若令 $Y=2y$, 则 $x^2+Y^2=9$. 这是新坐标 (x,Y) 的圆, 所以可以用前面的参数化: $x=3\cos(\theta)$ 和 $Y=3\sin(\theta)$, $0\leqslant\theta<2\pi$. 现在只需写出 $y=Y/2$ 来得到椭圆的参数化

$$x=3\cos(t)\quad 和 \quad y=\frac{3}{2}\sin(t), 其中\ 0\leqslant t<2\pi$$

当然, 这不是唯一的参数化!

那 $x^6+y^6=64$ 呢? 这个曲线留给你来画, 你可以看到该图像就像是一个膨胀的 "半径" 为 $64^{1/6}=2$ 单位的圆. 这启发了我们可以调整前面的圆的参数化. 首先, 我们需要将半径改为 2 单位: 其实, $x=2\cos(t)$ 和 $y=2\sin(t)$ 适合圆 $x^2+y^2=4$, 但不能参数化膨胀的圆, 因为 $\cos^6(t)+\sin^6(t)=1$ 一般不成立. 那怎么调整它呢? 我们可用 $\cos(t)$ 的某些次幂来替换它, 当对它们取 6 次方时, 可得 $\cos^2(t)$. 这应该是 $\cos^{1/3}(t)$. 所以, 若令 $x=2\cos^{1/3}(t)$ 和 $y=2\sin^{1/3}(t)$, 则应该可行. 我们来验证一下:

$$x^6+y^6=(2\cos^{1/3}(t))^6+(2\sin^{1/3}(t))^6=64\cos^2(t)+64\sin^2(t)=64,$$

这正是我们想要的. 为了得到整个曲线, 如前面一样, 我们令 t 在 0 到 2π 间取值.

参数方程的导数

这是一本微积分图书, 所以我们最好对这些参数求微积分. 要求曲线的切线方程, 当然要求导数. 由于 x 和 y 都是关于 t 的函数, 所以要用到链式法则. 就是说

$$\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t},$$

两边除以 $\mathrm{d}x/\mathrm{d}t$, 整理后得

$$\boxed{\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}.}$$

若把 x 看作 $x(t)$, y 类似, 则可将该方程另写为

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{y'(t)}{x'(t)}.$$

我们用 3 个例子来看一下如何应用.

首先, 求参数曲线上对应于 $t=1/2$ 点的切线斜率和切线方程, 参数曲线定义为

$$x=\mathrm{e}^{-2t},\quad y=\sin^{-1}(t),\quad -1<t<1.$$

求导, 我们有

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -2\mathrm{e}^{-2t} \quad \text{和} \quad \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{\sqrt{1-t^2}}.$$

由于我们只关心点 $t = 1/2$, 因此需要立即求 $t = 1/2$ 处的导数, 可得

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -2\mathrm{e}^{-1} = -\frac{2}{\mathrm{e}} \quad \text{和} \quad \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{\sqrt{1-1/4}} = \frac{2}{\sqrt{3}}.$$

故在 $t = 1/2$, 我们有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{2/\sqrt{3}}{-2/\mathrm{e}} = -\frac{\mathrm{e}}{\sqrt{3}}.$$

太好了! 我们已经求出了斜率. 那切线呢? 该直线过点 (x, y) 且斜率为 $\mathrm{d}y/\mathrm{d}x$. 斜率知道了, 但 x 和 y 呢? 将 $t = 1/2$ 代入原 x 和 y 的方程, 可知 $x = \mathrm{e}^{-2\cdot(1/2)} = 1/\mathrm{e}$, $y = \sin^{-1}(1/2) = \pi/6$. 所以切线方程为

$$y - \frac{\pi}{6} = -\frac{\mathrm{e}}{\sqrt{3}}\left(x - \frac{1}{\mathrm{e}}\right),$$

稍作化简后为

$$y = -\frac{\mathrm{e}}{\sqrt{3}}x + \frac{1}{\sqrt{3}} + \frac{\pi}{6}.$$



现在看一个更棘手的例子：求曲线 $x^6 + y^6 = 64$ 在点 $(-2^{5/6}, 2^{5/6})$ 的切线方程. (应当将该点代入原方程验证它是否在曲线上.) 该问题可通过隐函数求导来完成, 不过这里我们用前一节末的参数化 $x = 2\cos^{1/3}(t)$ 和 $y = 2\sin^{1/3}(t)$ 来完成, 这里 $0 \leqslant t < 2\pi$. 求导得到

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{2}{3}\cos^{-2/3}(t)\sin(t) \quad \text{和} \quad \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{2}{3}\sin^{-2/3}(t)\cos(t).$$

故由链式求导法则,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{\frac{2}{3}\sin^{-2/3}(t)\cos(t)}{-\frac{2}{3}\cos^{-2/3}(t)\sin(t)} = -\frac{\cos^{5/3}(t)}{\sin^{5/3}(t)}.$$

我们想知道在点 $(-2^{5/6}, 2^{5/6})$ 会发生什么. 令 $x = -2^{5/6}$, 由于 $x = 2\cos^{1/3}(t)$, 可知 $2\cos^{1/3}(t) = -2^{5/6}$, 所以 $\cos(t) = -1/\sqrt{2}$. 若对 y 采用相同讨论, 会发现 $\sin(t) = 1/\sqrt{2}$. 现在可以求出 t 了 —— 若想求 t, 你应该可知 $t = 3\pi/4$ 是 0 到 2π 之间的唯一解. 不过无论如何, 你都不必求 t, 信不信由你! 知道 $\sin(t)$ 和 $\cos(t)$ 的值就足够代入前面 $\mathrm{d}y/\mathrm{d}x$ 的表达式得出

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\cos^{5/3}(t)}{\sin^{5/3}(t)} = -\frac{(-1/\sqrt{2})^{5/3}}{(1/\sqrt{2})^{5/3}} = 1.$$

故我们已经求出了切线的斜率为 1. 我们知道它过点 $(x, y) = (-2^{5/6}, 2^{5/6})$ 且斜率

为 1, 所以它的切线方程为

$$y-2^{5/6}=1(x-(-2^{5/6})).$$

要确保你能理解为什么它可以化简为

$$y=x+2^{11/6}.$$

现在来看最棘手的例子 (至少从概念上讲是如此). 给定参数方程

$$x=4t^2-4 \text{ 和 } y=2t-2t^3,\ t \text{ 为所有实数.}$$

这些方程描述了 x-y 平面的一条曲线, 我们来求一下该曲线在原点的任意切线方程. 注意我说的是 "任意" 而不是 "某个". 这是有原因的! 我们来求一下原点对应的值. 在原点, x 和 y 的值都为 0, 故 $x=4(t^2-1)=0$ 和 $y=2(t-t^3)=0$. 第一个方程仅当 $t^2=1$ 时成立, 故 t 值为 ±1; 这两个值都满足第二个方程. 结论是曲线在 $t=1$ 和 $t=-1$ 时都过原点. 现在我们知道

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}=\frac{2-6t^2}{8t}=\frac{1}{4t}-\frac{3t}{4}.$$

当 $t=1$, 我们有 $\mathrm{d}y/\mathrm{d}x=-1/2$, 所以切线过原点且斜率为 $-1/2$. 因此对应的切线方程为 $y=-x/2$. 另一方面, 当 $t=-1$, 我们有 $\mathrm{d}y/\mathrm{d}x=1/2$, 此时的切线方程为 $y=x/2$. 下面通过曲线的图像来说明为什么会是这样. 取一些 t 值并计算出相应的 x 和 y 值, 填表如下.

t	-2	$-\frac{3}{2}$	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2
x	12	5	0	-3	-4	-3	0	5	12
y	12	$\frac{15}{4}$	0	$-\frac{3}{4}$	0	$\frac{3}{4}$	0	$-\frac{15}{4}$	-12

描出这些点并做合理猜测, 曲线的图像应该如图 27-2 所示.

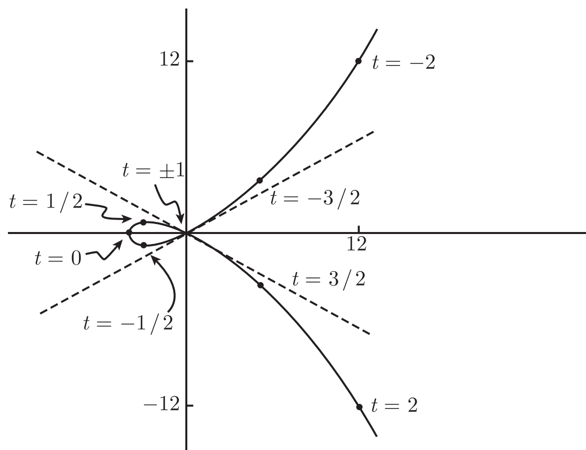


图 27-2

我们看到在原点确实有两条切线, 它们的斜率为 1/2 和 $-1/2$, 这看起来很合理.

假设我们求上述参数方程在 $t=1$ 处的二阶导数. 求 $\mathrm{d}^2y/\mathrm{d}x^2$ 的秘密就是将其看作 $\mathrm{d}y'/\mathrm{d}x$, 即把二阶导数看作 y' 的导数, 而 y' 是 y 关于 x 的导数. 现在问题变得简单了. 我们在前面已经看到

$$y'=\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{1}{4t}-\frac{3t}{4},$$

先不要将 $t=1$ 代入, 利用链式求导法则 (和 $x=4t^2-4$) 得

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}y'}{\mathrm{d}x}=\frac{\mathrm{d}y'/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}=\frac{\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{1}{4t}-\dfrac{3t}{4}\right)}{\dfrac{\mathrm{d}}{\mathrm{d}t}(4t^2-4)}=\frac{\dfrac{-1}{4t^2}-\dfrac{3}{4}}{8t}=-\frac{1}{32t^3}-\frac{3}{32t}.$$

最后代入 $t=1$ 得

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=-\frac{1}{32}-\frac{3}{32}=-\frac{1}{8}.$$

参照前面的图像做个检查. $t=1$ 对应的曲线部分是 y 轴左侧环线的上半部分, 它一直向下穿过原点移动到第四象限. 若你只关注曲线上原点附近的这部分, 可以看到这部分其实是下凹的, 故至少我们可以确信二阶导数为负, 这与前面发现的一样.

27.2 极 坐 标

假设你的朋友站在一块很大的平地上, 你俩都认为那是原点所在. 你会告诉他如何到达平地上的另一点. 如果利用笛卡儿坐标系, 你就会告诉你的朋友走到点 (x,y), 意思是你的朋友应该向东走 x 个单位, 然后再向北走 y 个单位. (你们得事先确定所用的单位.) 当然, 若 x 和 y 是负的, 意味着你朋友要往回走一定的单位. 你的朋友也可以先向北走 y 个单位, 然后再向东走 x 个单位, 仍然能到达相同的地点.

亦或, 你可以让你的朋友面朝东, 然后告诉他向逆时针方向转个角度 (还是站在原点). 若角度是负的, 意味着你的朋友是顺时针转. 然后, 让你的朋友沿着他面对着的方向行进一定的距离. 若该距离是负的, 则向相反方向行进. 此时与笛卡儿坐标 (x,y) 不同, 你朋友将到达 (r,θ), 这里 θ 是转过的角度, r 单位是行进的距离.

若你想描述的点是原点, 则可以告诉你朋友点为 $(0,\theta)$, θ 为任意角. 他转过多少角度是没有关系的, 因为没有行进, 所以他还是在原点. 还可以知道, 将 2π 加到角 θ 上没有差别, 他只是在 θ 基础上原地转一圈. 加 4π、6π 或其他任何 2π 的整数倍, 甚至负整数倍都是一样的, 加多少得看你有多狠心了, 让你的朋友在原地无目的地旋转很多圈只会让他眩晕! 现在来看看公式.

27.2.1 极坐标与笛卡儿坐标互换

考虑极坐标系下的点 (r,θ), 如图 27-3 所示.

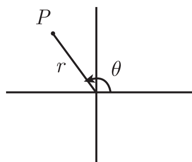


图 27-3

要知道, 你的朋友站在原点面朝 x 轴正方向, 然后逆时针旋转 θ 角, 然后向前行进 r 单位到达了点 P. 那么 P 的笛卡儿坐标 (x,y) 是什么呢? 我们知道 $\cos(\theta)=x/r$ 和 $\sin(\theta)=y/r$, 因此有

$$x=r\cos(\theta) \quad 和 \quad y=r\sin(\theta).$$

(将这个例子与 27.1 节的例子 $x=3\cos(t)$, $y=3\sin(t)$ 进行比较.) 总之, 这些方程展示了如何将极坐标转换为笛卡儿坐标. 例如, 极坐标系下的点 $(2, 11\pi/6)$ 的笛卡儿坐标是什么? 首先画图以便理解, 如图 27-4 所示.

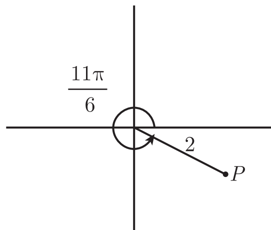


图 27-4

根据图可知相关角为 $2\pi-11\pi/6$, 即 $\pi/6$. 该点在第四象限, 所以余弦值为正且正弦值为负, 由此可得 $x=2\cos(11\pi/6)=2\cdot(\sqrt{3}/2)=\sqrt{3}$, 以及 $y=2\sin(11\pi/6)=2\cdot(-1/2)=-1$. 因此笛卡儿坐标是 $(\sqrt{3},-1)$.

通常将外语翻译成母语总比将母语翻译成外语容易些, 极坐标也一样. 从笛卡儿坐标转化到极坐标要稍难一些. 容易的部分是 r, 因为由勾股定理 (毕达哥拉斯定理) 知 $r^2=x^2+y^2$. (也可以通过将上面方框中的方程取平方, 然后加起来并运用 $\cos^2(x)+\sin^2(x)=1$ 来得到这个等式.) 那么 θ 呢? 我们知道若 $x\neq 0$, 则 $\tan(\theta)=y/x$, 但这并未告诉我们 θ 的确切值. 我们总可以将 π 的整数倍加到 θ 上而不改变 $\tan(\theta)$ 的值. 所以你应该画一个图看看具体情况. 上面情形可总结为:

$$r^2=x^2+y^2 \text{ 且 } \tan(\theta)=\frac{y}{x}, \ (x\neq 0), \text{但要检验象限!}$$

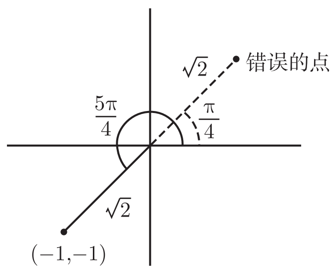


图 27-5

我们来看一个例子：假设要将 $(-1,-1)$ 写成极坐标的形式. 若将 $x=-1$ 和 $y=-1$ 代入上面的公式, 得到 $r^2=(-1)^2+(-1)^2=2$ 和 $\tan(\theta)=(-1)/(-1)=1$. 所以看似 $r=\sqrt{2}$ 和 $\theta=\tan^{-1}(1)=\pi/4$. 但其实是不对的! 检验图 27-5.

极坐标系下的点 $(\sqrt{2},\pi/4)$ 是错误的点, 因为它在第一象限. 正确的点在第三象限, 如你在图中所见, 极坐标应为 $(\sqrt{2},5\pi/4)$.

我们错在哪儿了呢? 实际上, 我们由 $\tan(\theta)=1$ 推出 $\theta=\pi/4$, 却忘了另一个答案 $\theta=5\pi/4$. 我们还得到 $r^2=2$, 所以 $r=\sqrt{2}$, 舍掉解 $r=-\sqrt{2}$. 若再看一下上面

的图, 可以看到点 $(-1,-1)$ 也可以写成极坐标 $(-\sqrt{2},\pi/4)$. 若你的朋友面朝错误的点站在原点, 然后往回走 $\sqrt{2}$ 个单位, 他最后也能走到正确的点处.

现在我们有两种方式将 $(-1,-1)$ 写成极坐标: $(\sqrt{2},5\pi/4)$ 和 $(-\sqrt{2},\pi/4)$. 但这还不全, 在 θ 上加 2π 的任何整数倍也是一样的. 故, 该点可用的所有极坐标如下:

$$\left(\sqrt{2},\frac{5\pi}{4}+2\pi n\right),\ \left(-\sqrt{2},\frac{\pi}{4}+2\pi n\right)\ (n\text{ 为整数}).$$

这表明有无穷多对 (r,θ) 在两个族中, 它们都描述了平面中的同一个点 $(-1,-1)$! 幸运的是, 几乎在每一个问题中只需要一对 (r,θ), 习惯上选择 $r\geqslant 0$ 且 θ 在 0 到 2π 之间的那对. 所以, 倘若你能理解这不是唯一的极坐标形式, 知道 $(-1,-1)$ 有极坐标 $(\sqrt{2},5\pi/4)$ 就可以了.

另外一些例子: 笛卡儿坐标为 $(0,1)$、$(-2,0)$ 和 $(0,-3)$ 的点的极坐标是什么? 在相同坐标轴下画出这些点, 如图 27-6 所示.

运用前面的公式 $\tan(\theta)=y/x$, 会遇到些麻烦. 例如, 在点 $(0,1)$ 处会得到 $\tan(\theta)=1/0$, 而这是没有定义的. 忘记这个公式吧! 只看图就能知道我们要找的角是 $\pi/2$, 故 $(0,1)$ 的极坐标为 $(1,\pi/2)$. 类似地, $(-2,0)$ 的极坐标为 $(2,\pi)$, $(0,-3)$ 的极坐标为 $(3,3\pi/2)$. 当然, 有无穷多个答案. 例如, 点 $(0,-3)$ 经常被写为极坐标 $(3,-\pi/2)$ 而不是 $(3,3\pi/2)$. 总之, 需要对很多点练习笛卡儿坐标和极坐标的互换, 直到熟练为止.

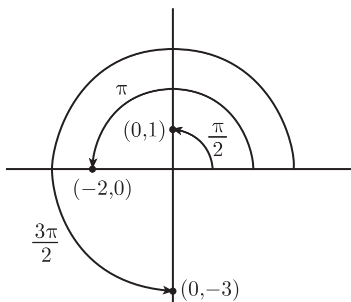


图 27-6

继续讲述之前, 让我们再想一想. 你记得在电影里看到的监视潜艇的发着绿光且发出 “哔、哔、哔 ……” 声音的雷达屏幕吗? 那些屏幕就如图 27-7 所示.

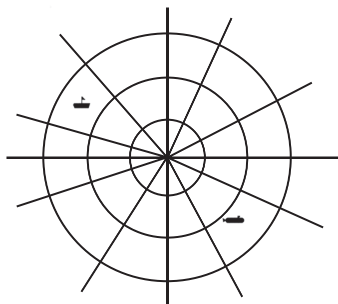


图 27-7

这就是极坐标系的 “格子”. 你知道, 一个格子包含 x 为常数的一些线 (垂直的线) 和 y 为常数的一些线 (水平的线). 若在极坐标系下, 则应该画出一些 r 为常数的曲线, 和一些 θ 为常数的曲线. r 等于某常数 C 的这些点构成了一个以原点为中心、C 个单位为半径的圆, 而 θ 为常数的点构成了一条以原点为起点的射线. 上图显示了一些这样的圆和射线. 你之前已经见过极坐标, 只不过从未意识到!

27.2.2 极坐标系中画曲线

假设某函数 f 的极坐标方程为 $r=f(\theta)$, 我们想在极坐标系下画出所有点 (r,θ)



的图像, 其中 $r = f(\theta)$ 在 θ 给定的范围内取值. 这个做起来不容易, 最好的方法也许是做个函数值表并描点. 先画出 $r = f(\theta)$ 在笛卡儿坐标系下的图像是有帮助的. 例如, 要画极坐标系下 $r = 3\sin(\theta)$ 的图像, 其中 $0 \leqslant \theta \leqslant \pi$, 我们先画出以 r 和 θ 为笛卡儿坐标轴的 $r = 3\sin(\theta)$ 图像, 见图 27-8.

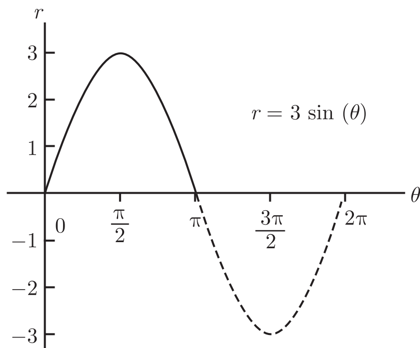


图 27-8

该图说明随着角由 0 转到 π, 距离 r 由 0 增加到 3, 然后在角度到 π 时回到 0. 所以所求的曲线如图 27-9 所示.

这个图看上去有点差劲. 为了更紧密些, 我们可以写出如下的函数值表.

θ	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	$2\pi/3$	$3\pi/4$	$5\pi/6$	π
r	0	$3/2$	$3/\sqrt{2}$	$3\sqrt{3}/2$	3	$3\sqrt{3}/2$	$3/\sqrt{2}$	$3/2$	0

画出这些点, 可得到图 27-10.

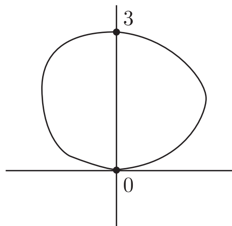


图 27-9

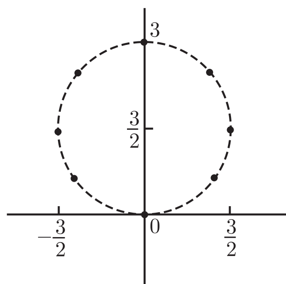


图 27-10

所以它其实像一个真正的圆, 不是那个看上去很蹩脚的图. 其实, 将其转换成笛卡儿坐标就可验证. 事实上, 由于 $y = r\sin(\theta)$, 且有 $r = 3\sin(\theta)$, 因而可以消除 θ 得到 $r^2 = 3y$. 另一方面, $r^2 = x^2 + y^2$, 故我们得到 $x^2 + y^2 = 3y$. 将 $3y$ 移到左边并凑成关于 y 的完全平方式可得 $x^2 + (y - 3/2)^2 = (3/2)^2$, 它是以 $(0, 3/2)$ 为圆心、3/2 为半径的圆. 这与上面的图像一致. 现在, 可自行验证若 θ 从 π 变到 2π, 结果只是沿着该圆又描一遍.





我们来看另一个例子. 假定要画曲线 $r = 1 + 2\cos(\theta)$, 其中 $0 \leqslant \theta \leqslant 2\pi$ 的图像. 首先, 注意笛卡儿图像 (见图 27-11).

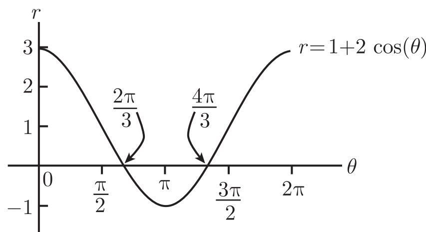


图 27-11

求出图像在 θ 轴 (就是通常被认为是 x 轴的水平轴) 的交点是很重要的. 你知道, 在 θ 轴的交点处 $r = 0$, 所以这时极坐标系下的图像应该回到原点. 在这个例子中, 我们有 $1 + 2\cos(\theta) = 0$, 意味着 $\cos(\theta) = -1/2$. 由于 $\cos(\theta)$ 是负的, 因此 θ 必须在第二或第三象限. 而且相应的角度是 $\cos^{-1}(1/2)$, 即 $\pi/3$. 我们推出了当 $\theta = 2\pi/3$ 或 $4\pi/3$ 时 $r = 0$, 如图 27-11 所示.

现在画 $r = 1 + 2\cos(\theta)$ 的极坐标图像. 随着 θ 从 0 增加到 $2\pi/3$, 距离 r 从 3 递减到 0, 当 $\theta = \pi/2$ 时穿过 1. 这是目前为止我们可得到的图, 如图 27-12 所示.

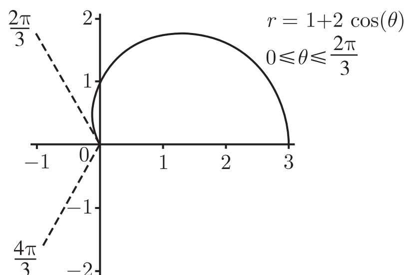


图 27-12

随着 θ 从 $2\pi/3$ 增加到 π, 距离 r 递减到 -1. 这意味着我们要折回到第四象限, 而不是停留在第二象限, 图 27-13 可以说明.

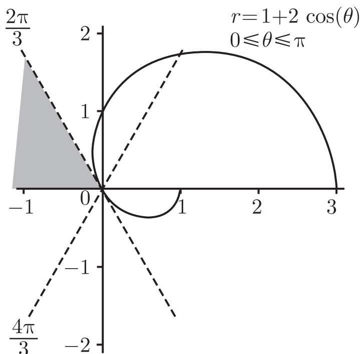


图 27-13

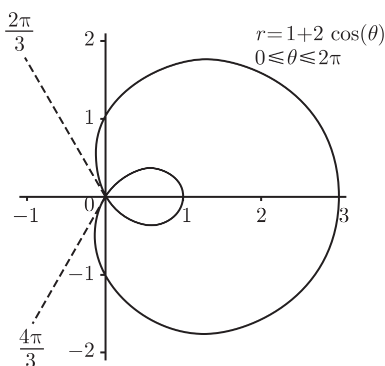
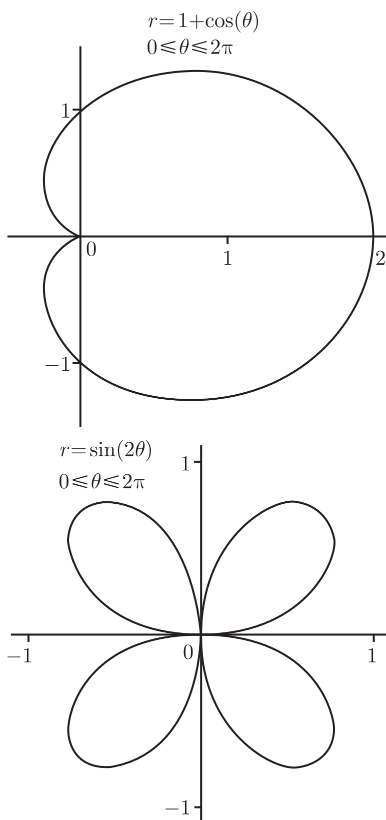


图 27-14

随着 θ 从 $2\pi/3$ 增加到 π, 图像应该包含在阴影区域, 但由于这时的 r 是负的, 所以图像一下转到了第四象限. 不管怎样, 我们可以用这种方法继续讨论直到 $\theta = 2\pi$, 或者至少注意到 $r = 1 + 2\cos(\theta)$ 的笛卡儿图像关于直线 $\theta = \pi$ 是对称的, 这意味着我们要画的完整图像就是现有图像 (关于水平轴) 的镜面映像 (见图 27-14).

最后, 我们来观察一些选定的极坐标曲线 (见图 27-15). 你可能想全部拿下这些图像, 并试图画出图来, 或者说服自己相信每个图都是正确的. 不管怎样, 你应该画很多极坐标曲线, 直到自己感觉不再有进展了.



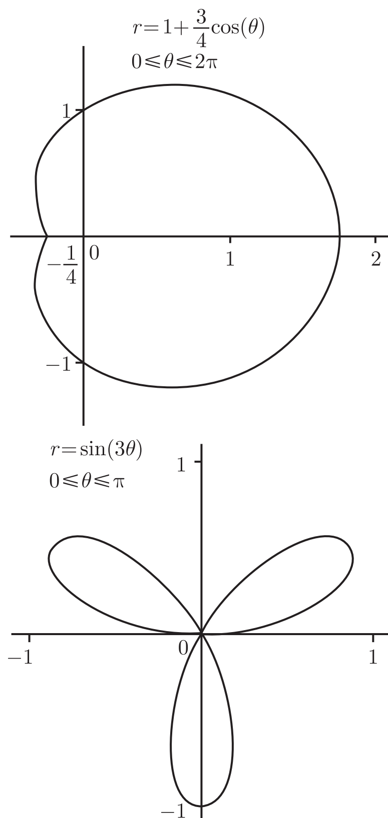


图 27-15

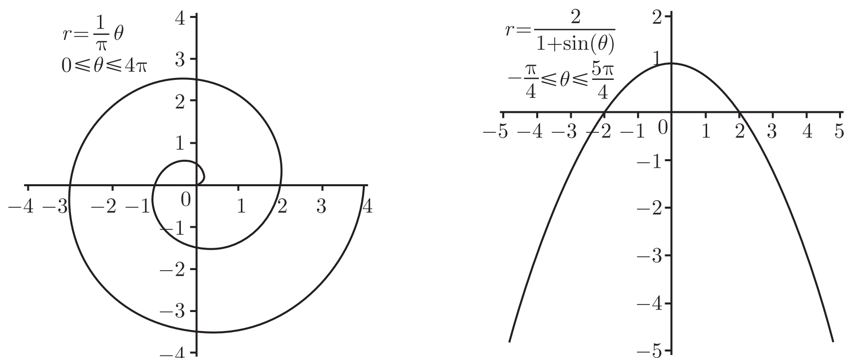


图　27-15　(续)

关于上述曲线的一些事实.

(1) 由 $r = 1 + \cos(\theta)$ 确定的曲线称为心形线. 曲线 $r = 1 + \frac{3}{4}\cos(\theta)$ 是蜗牛形曲线的一个例子, 而心形线是蜗牛形曲线的特例.

(2) 在 $r = \sin(3\theta)$ 的图像中, 角 θ 只从 0 取到 π. 当 θ 从 π 取到 2π 时, 图像折了回来, 就像圆 $r = \sin(\theta)$ 的情形一样.

(3) 曲线 $r = \theta / \pi$ 是阿基米德螺旋线的一个例子, 该曲线不是周期的：随着 θ 的增加, 螺旋变得越来越大.



(4) 曲线 $r = 2/(1 + \sin(\theta))$ 看起来像一个抛物线. 实际上, 可以证明给定的方程在笛卡儿坐标下为 $x^2 = 4 - 4y$.

27.2.3　求极坐标曲线的切线

幸运的是, 求极坐标曲线的切线就是求参数方程确定的曲线的切线的特殊情形. 我们已经在第 27 章开始讨论过这一问题的一般求解方法. 我们来看一下在极坐标下怎么运用这个方法.

我们有 $r = f(\theta)$, 要求该曲线上某点处的切线. 利用 $x = r\cos(\theta)$ 和 $y = r\sin(\theta)$, 我们有

$$x = f(\theta)\cos(\theta) \quad \text{和} \quad y = f(\theta)\sin(\theta),$$

这意味着 x 和 y 都被 θ 参数化了. 根据第 27 章开始部分的公式, 我们有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta},$$

这给出了通常的切线斜率. 最后, 只需代入我们所关心的 θ. 这就是该问题的全部, 来看一下实例.



考虑极坐标下的曲线 $r = 1 + 2\cos(\theta)$. 我们在前一节画过该曲线图像, 假设我们要求穿过极坐标为 $(2, \pi/3)$ 的点的切线方程. 首先检查一下这个点在曲线上吗？当 $\theta = \pi/3$, 有 $1 + 2\cos(\theta) = 1 + 2\cos(\pi/3) = 2$, 即给出了 r 的值. 所以该

点的确在曲线上. 下一步, 我们要求出切线的斜率 $\mathrm{d}y/\mathrm{d}x$. 我们有 $x = r\cos(\theta) = (1+2\cos(\theta))\cos(\theta)$ 和 $y = r\sin(\theta) = (1+2\cos(\theta))\sin(\theta)$, 需要求出 $\mathrm{d}y/\mathrm{d}\theta$ 和 $\mathrm{d}x/\mathrm{d}\theta$. 不幸的是, 这里要用到积的求导法则, 不过还不算太糟, 自行验证



$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = -2\sin^2(\theta) + (1+2\cos(\theta))\cos(\theta) \text{ 和 } \frac{\mathrm{d}x}{\mathrm{d}\theta} = -\sin(\theta)(1+4\cos(\theta)).$$

所以, 我们有

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\theta}{\mathrm{d}x/\mathrm{d}\theta} = \frac{-2\sin^2(\theta) + (1+2\cos(\theta))\cos(\theta)}{-\sin(\theta)(1+4\cos(\theta))}.$$

我们想知道当 $\theta = \pi/3$ 时会怎样. 将其代入, 应该可得

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-2(3/4) + (1+2(1/2))(1/2)}{-(\sqrt{3}/2)(1+4(1/2))} = \frac{1}{3\sqrt{3}}.$$

我们知道了所求直线的斜率. 现在只需找到直线穿过的一个点. 显然那个点的极坐标为 $(2, \pi/3)$, 不过我们需要它的笛卡儿坐标. 因此, 只需用 $x = r\cos(\theta)$ 和 $y = r\sin(\theta)$ 来得到 $x = 2\cos(\pi/3) = 1$ 和 $y = 2\sin(\pi/3) = \sqrt{3}$. 太好了, 我们要求的直线过点 $(1, \sqrt{3})$, 且斜率为 $1/(3\sqrt{3})$. 这条线为

$$y - \sqrt{3} = \frac{1}{3\sqrt{3}}(x-1),$$

稍加化简后得到答案

$$y = \frac{1}{3\sqrt{3}}(x+8),$$

那这条曲线在原点处的切线呢? 见 27.2.2 节中 $r = 1 + 2\cos(\theta)$ 的图像, 可以看到在那点应该有两条切线! 不过, 我们还是能求出它们的方程. 事实上, 我们知道当 $r = 0$ 时曲线过原点, 且在前一节看到此时 $\theta = 2\pi/3$ 或 $\theta = 4\pi/3$. 可以验证, 将这些 θ 值分别代入前面 $\mathrm{d}y/\mathrm{d}x$ 的方程可得到 $-\sqrt{3}$ 和 $\sqrt{3}$. 由于两条切线过原点, 它们必有方程 $y = -\sqrt{3}x$ 和 $y = \sqrt{3}x$. 事实上, 这些直线补齐了对应于 $\theta = 2\pi/3$ 或 $\theta = 4\pi/3$ 的射线, 如 27.2.2 节图 27-13 上的虚线所示.



27.2.4 求极坐标曲线围成的面积

若想求由极坐标曲线 $r = f(\theta)$ 围成的面积, 其中 f 假设是连续的, 则需要求积分. 接下来呢? 我们只需建立正确的黎曼和. (回顾黎曼和, 参见 16.2 节.) 假设我们取介于 θ 和 $\theta + \mathrm{d}\theta$ 间的一小块角. 这块角沿逆时针移动, r 则从 $f(\theta)$ 缓慢移动到 $f(\theta + \mathrm{d}\theta)$. 若 $\mathrm{d}\theta$ 很小, 则 r 不会距离 $f(\theta)$ 很远, 所以可以用半径为 $r = f(\theta)$、角为 $\mathrm{d}\theta$, 以原点为中心的一小块饼图来近似所求的楔形, 如图 27-16 所示.

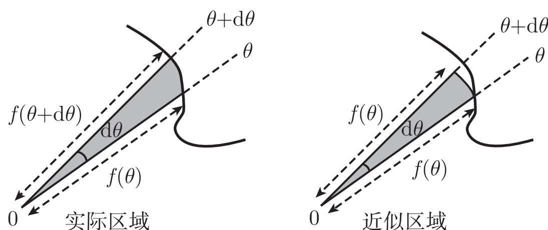


图　27-16

扇形的面积是半径平方的二分之一乘以扇形的角 (当然是弧度角). 所以, 可以近似楔形的面积为 $\frac{1}{2}(f(\theta))^2\mathrm{d}\theta$ (平方单位), 即 $\frac{1}{2}r^2\mathrm{d}\theta$. 当 θ 从 θ_0 变到 θ_1 时, 整个面积可通过将所有的楔形面积加起来, 并令 $\mathrm{d}\theta$ 递减于 0 来得到. 即有[①]下面的积分：

$$(\text{在 } r=f(\theta) \text{ 之内, 介于 } \theta=\theta_0 \text{ 和 } \theta=\theta_1 \text{ 之间的面积})=\int_{\theta_0}^{\theta_1}\frac{1}{2}r^2\mathrm{d}\theta.$$

跟前面一样, 面积以平方为单位.

我们将这个公式用于曲线 $r=3\sin(\theta)$, 其中 $0\leqslant\theta\leqslant\pi$. 由 27.2.2 节可知, 该曲线是一个半径为 3/2 个单位的圆, 所以它的面积应该为 $\pi(3/2)^2$, 即 $9\pi/4$ 平方单位. 我们来证明它. 我们有

$$\text{面积}=\int_0^{\pi}\frac{1}{2}r^2\mathrm{d}\theta=\frac{1}{2}\int_0^{\pi}(3\sin(\theta))^2\mathrm{d}\theta=\frac{9}{2}\int_0^{\pi}\sin^2(\theta)\mathrm{d}\theta.$$

这个积分可以用二倍角公式来求解, 就像 19.1 节开始描述的那样. 请验证答案为 $9\pi/4$.

这是一个更难的例子. 我们求由曲线 $r=1+2\cos(\theta)$ 围成的形如新月形面包的区域的面积, 如图 27-17 所示.

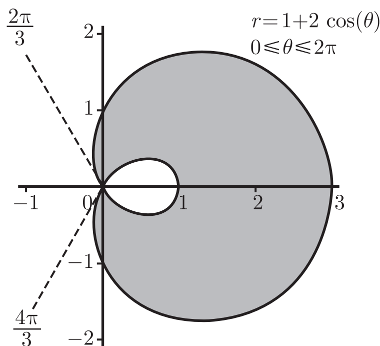


图　27-17

① 要证明这个公式, 需要考虑 $f(\theta)$ 的最大值和最小值来确定面积的上和与下和, 其中 θ 取值于 $[\theta_0,\theta_1]$ 的子区间, 然后证明当分割的区间趋于 0 时, 上和与下和都收敛于相同的值.

看来应该能够利用公式求得我们要求的面积为

$$\int_0^{2\pi} \frac{1}{2} r^2 \mathrm{d}\theta = \frac{1}{2} \int_0^{2\pi} (1 + 2\cos(\theta))^2 \mathrm{d}\theta.$$



同样, 求这个积分还要用到二倍角公式. 你可自己证明

$$\frac{1}{2} \int (1 + 2\cos(\theta))^2 \mathrm{d}\theta = \frac{3}{2}\theta + 2\sin(\theta) + \frac{1}{2}\sin(2\theta) + C.$$

上面的定积分可以通过代入 $\theta = 2\pi$ 和 $\theta = 0$ 并相减 (即 3π) 而求得. 不幸的是, 这不是正确的答案. 问题出在当 θ 位于 $2\pi/3$ 和 $4\pi/3$ 之间时, r 为负. 由于面积公式中包含 r^2, 因此无法辨别正负面积. (这与笛卡儿坐标下的情况大不相同, 在笛卡儿坐标系中, y 轴以下都为负.) 所以, 我们刚才求得的是在曲线 $r = |1 + \cos(2\theta)|$ 里的面积, 如图 27-18 所示.

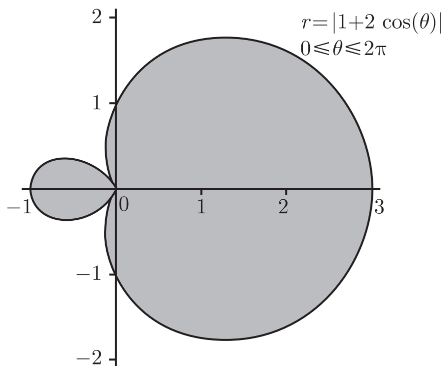


图 27-18

为了修正这个情况, 我们需要求竖轴左侧的小环内的面积, 然后从原面积中减去两次. 为什么是两次? 因为减去一次只是得到上图剩下的阴影部分面积, 而我们想求的面积其实是在该区域再剪去一个小环. 那么如何求小环内的面积呢? 重复上面的积分, 不过这次从 $2\pi/3$ 到 $4\pi/3$:

$$\text{小环的面积} = \frac{1}{2} \int_{2\pi/3}^{4\pi/3} (1 + 2\cos(\theta))^2 \mathrm{d}\theta.$$



现在用前面的反导来求积分值, 结果为 $(\pi - 3\sqrt{3}/2)$ 平方单位. 因此, 最后可以将我们要求的面积表示成 3π 平方单位减去两倍的小环面积, 然后计算出该面积:

$$\text{我们要求的面积} = 3\pi - 2\left(\pi - \frac{3\sqrt{3}}{2}\right) = (\pi + 3\sqrt{3})\ \text{平方单位}.$$

如该例所示, 若 r 可能为负, 在利用上述公式求极坐标系下的面积时要非常小心.

第 28 章　复　　数

为什么有些二次方程很有趣呢？二次方程 $x^2-1=0$ 可以有两个根 (1 和 -1)，而可怜的 $x^2+1=0$ 则没有根，因为它的判别式为负. 为了平衡，我们引入了复数的概念. 利用复数，任何一个二次方程都有两个根.[①] (可认为 $(x-a)^2=0$ 的复根 a 为两个根.) 下面是我们即将讨论的复数内容：

- 基本运算 (加法、减法、乘法、除法) 及解二次方程;
- 复平面及复数的笛卡儿和极坐标形式;
- 复数的高次幂;
- 解形如 $z^n=w$ 的方程;
- 解形如 $e^z=w$ 的方程;
- 利用幂级数和复数的一些技巧求解一些级数问题.

28.1　基　　础

不能取 -1 的平方根着实有点令人失望. 然而，我们接下来就要做这件事. 我们创造一个 -1 的平方根，称之为 i. 这样我们就有了 $\mathrm{i}^2=-1$. i 是 -1 唯一的平方根吗？不，如果这个世界是公平的，$-\mathrm{i}$ 也应该是一个平方根，则

$$(-\mathrm{i})^2=(-1)^2(\mathrm{i})^2=1(-1)=-1.$$

(事实上，世界是公平的，这一系列等式是正确的.) 由于 $\mathrm{i}^2+1=0$ 且 $(-\mathrm{i})^2+1=0$，则二次方程 $x^2+1=0$ 有了两个根：但它们不是实根，它们是虚根. 那 2i 呢？它也是虚的. 实际上，$(2\mathrm{i})^2=2^2\mathrm{i}^2=4(-1)=-4$，所以 $(2\mathrm{i})^2$ 是一个负数. 故，我们说一个数是虚数，意思是它的平方是一个负数. 虚数的唯一形式为 $y\mathrm{i}$，其中 y 是不等于 0 的实数. 也可用 $\mathrm{i}y$ 代替 $y\mathrm{i}$ 表示虚数.

现在，你可以对实数和复数进行加减了，如 $2-3\mathrm{i}$，但结果不能化简. 用这种方法，我们得到所有的复数，即所有形为 $x+\mathrm{i}y$ 的数，其中 x 和 y 为实数. 全体复数的集合通常用符号 $\mathbb{C}$ 来表示. 注意，所有虚数都是复数，如 $2\mathrm{i}=0+2\mathrm{i}$；所有实数也都是复数，如 $-13=-13+0\mathrm{i}$. 每个复数都有实部和虚部. 若 $z=x+\mathrm{i}y$，则实部是 x，虚部是 y，分别被写作 $\mathrm{Re}(z)$ 和 $\mathrm{Im}(z)$. 例如，$\mathrm{Re}(2-3\mathrm{i})=2$ 和 $\mathrm{Im}(2-3\mathrm{i})=-3$. 注意，$\mathrm{Im}(2-3\mathrm{i})$ 不是 $-3\mathrm{i}$ 而是 -3. $\mathrm{Re}(2\mathrm{i})$ 是什么呢？将 2i 写成 $0+2\mathrm{i}$，可以看到

---

① 令人惊奇的是，它对更高次的多项式也成立：任何次数为 n 的多项式有 n 个复数根 (重根也算在内). 这是源于所谓的代数基本定理，但那个方法不在本书讨论范围内. 或许可在关于复分析的书中找到更多相关内容.

实部是 0. 另一方面, 虚部 Im(2i) 显然是 2.

复数的加减法很简单. 就是将实部相加 (或相减), 然后再处理虚部. 例如,

$$(2-3\mathrm{i})+(-6-7\mathrm{i})=2-6-3\mathrm{i}-7\mathrm{i}=-4-10\mathrm{i};$$

减法的例子是

$$(2-3\mathrm{i})-(-6-7\mathrm{i})=2+6-3\mathrm{i}+7\mathrm{i}=8+4\mathrm{i}$$

乘法也不难, 只需展开, 但要记住每次见到 i^2 时都要把它换成 -1. 例如,

$$\begin{aligned}(2-3\mathrm{i})(-6-7\mathrm{i})&=2(-6)+2(-7\mathrm{i})-(3\mathrm{i})(-6)-(3\mathrm{i})(-7\mathrm{i})\\&=-12-14\mathrm{i}+18\mathrm{i}+21\mathrm{i}^2=-12+4\mathrm{i}-21=-33+4\mathrm{i}.\end{aligned}$$

那 i^3 是什么呢？i^4 呢？i^5 呢？我们从 i^3 开始. 我们有 $\mathrm{i}^3=\mathrm{i}^2\times\mathrm{i}=(-1)\times\mathrm{i}=-\mathrm{i}$, 所以 i^3 就是 $-\mathrm{i}$. $\mathrm{i}^4=\mathrm{i}^3\times\mathrm{i}=(-\mathrm{i})\times\mathrm{i}=1$, 即 $\mathrm{i}^4=1$. 对于 i^5, 用相同的方法：$\mathrm{i}^5=\mathrm{i}^4\times\mathrm{i}=1\times\mathrm{i}=\mathrm{i}$. 事实上, 由于 $\mathrm{i}^4=1$, 我们可以看到 i 的幂次在 1、i、-1、$-\mathrm{i}$ 中循环. 例如, $\mathrm{i}^{101}=\mathrm{i}$, 因为 $\mathrm{i}^{100}=1$ (要知道 100 可被 4 整除).

除法呢？有点棘手, 不过还好. 方法与分母有理化很相似. 该方法源于如下的观察：如果有一个复数 $x+\mathrm{i}y$, 并令其乘以复数 $x-\mathrm{i}y$, 得到一个实数. 当做计算时, 会意识到应用平方差公式：

$$(x+\mathrm{i}y)(x-\mathrm{i}y)=x^2-(\mathrm{i}y)^2=x^2-\mathrm{i}^2y^2=x^2+y^2.$$

x 和 y 都是实数, 显然 x^2 和 y^2 也是, 故它们的和也是实数. 若 $z=x+\mathrm{i}y$, 与其对应的 $x-\mathrm{i}y$ 是如此重要, 故它有个名字：**共轭复数**, 并表示为 $\bar{z}$. 例如, 若 $z=2-3\mathrm{i}$, 则 $\bar{z}=2+3\mathrm{i}$; 若 $z=7\mathrm{i}$, 则 $\bar{z}=-7\mathrm{i}$. 注意, 实数的共轭复数仍是该实数, 因为在取共轭复数时, 只是变换了虚部的符号, 但实数的虚部为 0. 如前面的公式所示, 一个数与它的共轭复数相乘得实数, 即实部和虚部的平方和. 受勾股定理和上面的公式启发, 对给定的复数 $z=x+\mathrm{i}y$, 我们定义 z 的模为 $\sqrt{x^2+y^2}$. 将 z 的模写作 $|z|$, 则

$$|x+\mathrm{i}y|=\sqrt{x^2+y^2}.$$

这里是一些例子：$|2-3\mathrm{i}|=\sqrt{2^2+(-3)^2}=\sqrt{4+9}=\sqrt{13}$. 类似地, $|7\mathrm{i}|=\sqrt{0^2+7^2}=7$. 那 $|-13|$ 呢？我们有 $|-13|=\sqrt{(-13)^2+0^2}=13$, 即 -13 的绝对值. 模的表示符号与原来绝对值的表示符号完全一致. 其实, 可认为模是绝对值的加强版. 不管怎样, 前面的平方差公式显示了复数与它的共轭复数的乘积是模的平方, 即

$$z\bar{z}=|z|^2.$$

完成这些准备工作之后, 我们该讨论复数除法了. 你要做的就是上下同乘分母部分的共轭复数, 然后展开. 新的分母是原分母模的平方, 如

$$\frac{2-3\mathrm{i}}{-6-7\mathrm{i}}=\frac{(2-3\mathrm{i})(-6+7\mathrm{i})}{(-6-7\mathrm{i})(-6+7\mathrm{i})}.$$

分子部分需要完全展开, 分母就是 $|-6-7\mathrm{i}|^2$, 所以

$$\frac{2-3\mathrm{i}}{-6-7\mathrm{i}}=\frac{-12+18\mathrm{i}+14\mathrm{i}-21\mathrm{i}^2}{(-6)^2+(-7)^2}=\frac{9+32\mathrm{i}}{85}=\frac{9}{85}+\frac{32}{85}\mathrm{i}.$$

我们可推出

$$\mathrm{Re}\left(\frac{2-3\mathrm{i}}{-6-7\mathrm{i}}\right)=\frac{9}{85}\quad 和\quad \mathrm{Im}\left(\frac{2-3\mathrm{i}}{-6-7\mathrm{i}}\right)=\frac{32}{85}.$$

另一个例子：求

$$\mathrm{Re}\left(\frac{3+4\mathrm{i}}{\mathrm{i}-1}\right).$$

这个例子有个小陷阱. 分母其实应该写成 $-1+\mathrm{i}$. 这样做了, 就能看到分母的共轭复数是 $-1-\mathrm{i}$, 所以

$$\frac{3+4\mathrm{i}}{\mathrm{i}-1}=\frac{(3+4\mathrm{i})(-1-\mathrm{i})}{(-1+\mathrm{i})(-1-\mathrm{i})}=\frac{-3-3\mathrm{i}-4\mathrm{i}-4\mathrm{i}^2}{(-1)^2+(1)^2}=\frac{1-7\mathrm{i}}{2}=\frac{1}{2}-\frac{7}{2}\mathrm{i}.$$

所以 $(3+4\mathrm{i})/(\mathrm{i}-1)$ 的实部为 $\frac{1}{2}$, 它的虚部为 $-\frac{7}{2}$.

我们来看看如何解二次方程. 例如, 欲解 $x^2+3x+14=0$, 只需用二次方程公式和 $\sqrt{-1}=\mathrm{i}$ 来得出

$$x=\frac{-3\pm\sqrt{3^2-4\times1\times14}}{2}=\frac{-3\pm\sqrt{-47}}{2}=-\frac{3}{2}\pm\frac{\sqrt{47}}{2}\mathrm{i}.$$

注意, 我们已将 $\pm\sqrt{-47}$ 化简为 $\pm\sqrt{47}\cdot\mathrm{i}$. 如果是系数为复数的二次方程呢？二次方程公式仍可用, 但可能要求复数的平方根, 而不只是刚刚做的只是求负数的平方根. 我们将在 28.4.1 节看到这样的例子.

复指数函数

我们已经讨论了如何加、乘复数. 那么, 如何指数化它们呢？我们来看如何使形如 e^z 的数有意义, 其中 z 是复数. 从 24.2.3 节知

$$\mathrm{e}^x=\sum_{n=0}^{\infty}\frac{x^n}{n!},$$

对所有实数 x 都成立. 如果我们将右边的 x 换成 z(其中 z 为复数) 会发生什么？我们将得到一个项为复数的级数. 不管相信与否, 你仍可用比式判别法证明该级数收敛, 无论 z 是什么样的复数. (我们只证明了实数级数的比式判别法, 但你一旦定义了复数序列的收敛, 该证明仍成立.) 受所有这些启发, 我们对任意复数 z, 通过等式

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

定义 e^z. 该等式当 z 为实数时当然成立, 因为 e^x 满足上面等式. 另一方面, 如果新对象 e^z 能满足我们对指数的所有预期就好了. 其实, 关键是满足指数法则 $e^z e^w = e^{z+w}$. 一旦我们知道了这个, 其他所有的指数法则马上也能多少得到满足.

那么, 怎么证明 $e^z e^w = e^{z+w}$? 这里有一个间接的方法. 我们知道 $e^x e^y = e^{x+y}$ 对任意实数 x 和 y 成立, 这意味着

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{m=0}^{\infty} \frac{y^m}{m!} = \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!}.$$

我们只是将每个指数用它们的麦克劳林级数代换了, 在每个和中用了不同的虚拟变量. 如果将左边的两个级数乘开, 将得到一些 x 和 y 幂次的双幂级数; 右边同理. 因此, 等式左边和右边 $x^n y^m$ 的系数相同, 若将 x 和 y 用复数 z 和 w 分别代替, 同样也成立. 因此, 我们证明了 $e^z e^w = e^{z+w}$ 对任意两个复数 z 和 w 成立!

28.2 复 平 面

实数常常被表示为数轴上的点, 是一维的. 从字面上看, 复数还多一维. 其实, 若 $z = x + iy$, 我们不能将所有信息压缩到一个实数上去. 我们将采用复平面而非实数轴. 复数 $z = x + iy$ 将用笛卡儿坐标系下的 (x, y) 表示. 画形如 $2 - 3i$、$2i$ 和 -1 的复数是很简单的, 如图 28-1 所示.

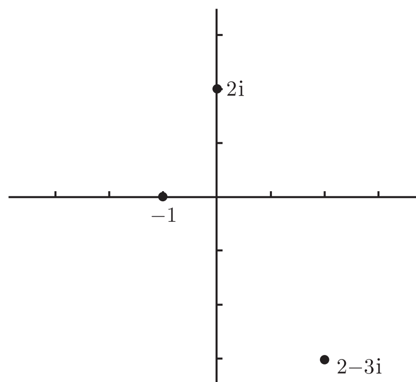


图 28-1

你应该将每个点看作一个复数, 而不是一对实数.

在前一章我们看到, 可将平面中的点用极坐标代替. (如果很久没看了, 现在应该复习 27.2.1 节.) 那么如果你有复平面内极坐标为 (r, θ) 的点, 该点所表示的复数是什么呢? 我们可用 $x = r\cos(\theta)$ 和 $y = r\sin(\theta)$ 来转化到笛卡儿坐标. 所以极坐标 (r, θ) 表示复数 $z = x + iy = r\cos(\theta) + ir\sin(\theta)$. 特别的, 若 $r = 1$, 则 z 就是 $\cos(\theta) + i\sin(\theta)$.

欧拉给出了一个奇异且独特的等式, 它很重要:

$$\boxed{e^{i\theta} = \cos(\theta) + i\sin(\theta).}$$

对所有实数 θ 都成立.[①] 意思是, 按前一节定义复数 $e^{i\theta}$, 当在复平面上画该点时有

① 该等式证明见本章末.

极坐标 $(1,\theta)$. 所以 $e^{i\theta}$ 在单位圆上且有从 x 轴正方向开始的角 θ. 图 28-2 给出了不同 θ 值对应的 $e^{i\theta}$.

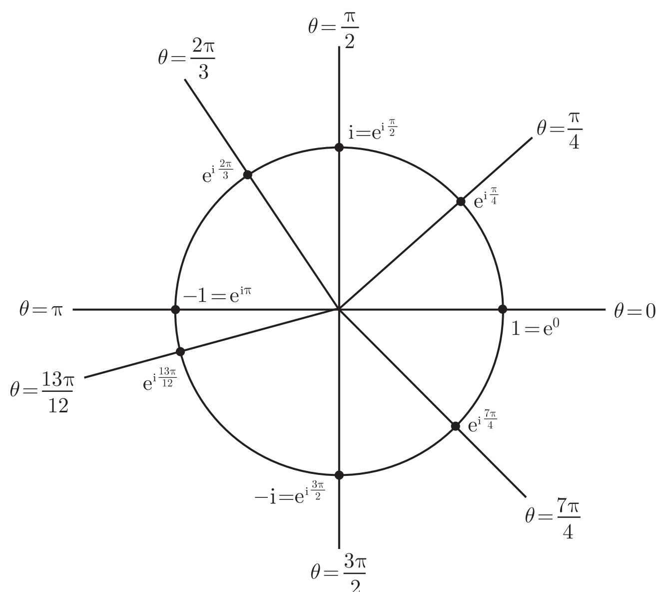


图　28-2

对于不在单位圆上的点, 你只需乘以 r. 特别地, 我们看到如果 z 由极坐标系下的点 (r,θ) 表示, 则 $z = r\cos(\theta) + ir\sin(\theta)$. 由欧拉等式, 这意味着 $z = re^{i\theta}$, 故我们证得

若 (x,y) 和 (r,θ) 为相同的点, 则 $x + iy = re^{i\theta}$.

我们说, 形如 $re^{i\theta}$ 的复数为极坐标形式 (与之相对的 $x+iy$ 为笛卡儿形式). 例如, 在上图中 $-1 = e^{i\pi}$, 这是因为笛卡儿坐标点 $(-1,0)$ 有极坐标 $(1,\pi)$, 所以 $-1+0i = 1e^{i\pi}$. 也就是说, -1 的极坐标形式是 $e^{i\pi}$. 类似地, 笛卡儿坐标点 $(0,1)$ 可以写成极坐标 $(1,\pi/2)$, 所以 $0 + 1i = 1e^{i\pi/2}$, 或 $i = e^{i\pi/2}$. 这个式子看起来有点奇怪, 但确实是正确的：左边是笛卡儿形式, 而右边是极坐标形式. 相同的还有 $-i = e^{i(3\pi/2)}$. (知道为什么吗？)

在 27.2.1 节, 我们看到有无穷多种方式来表示一个给定的极坐标点. 当我们讨论复数时与之一致, 这里令 r 非负. 同样地, 如果已经求出给定点的极坐标 (r,θ), 则可以将 2π 的任意整数倍加到 θ 上, 结果不变. 例如, 点 $(0,-1)$ 有极坐标 $(1,3\pi/2)$, 或者减去 2π 得到该点的另一个极坐标 $(1,-\pi/2)$. 对于复数, 这意味着 $e^{i(3\pi/2)} = e^{-i\pi/2}$. 所以 $e^{i\theta}$ 关于 θ 是**周期的**, 且周期为 2π. 这个结果很重要, 稍后将用到.

前面讨论了 $e^{i\pi} = -1$, 我们仔细想想, 也许你会觉得, 这真是太令人惊叹了. 在

你的数学学习中, 有多少个基本新数呢? 引入 -1 打开了通往负数的大门. 数字 π 来自圆的几何. 等字 e 是自然对数的底, 在微积分学习中很重要. 数字 i 指引我们通往复数的路并得以求解二次 (和更高次多项式) 方程. 如果你问我, 我会说它们结合成这样简单的公式真是很不寻常. 好了, 哲学闲谈就到此, 我们来看一些复数的极坐标形式与笛卡儿形式相互转换的例子.

笛卡儿形式和极坐标形式互换

将极坐标形式的复数转换成笛卡儿形式, 可以直接应用欧拉恒等式, 即 $\mathrm{e}^{\mathrm{i}\theta} = \cos(\theta)+\mathrm{i}\sin(\theta)$. 例如, $2\mathrm{e}^{\mathrm{i}(5\pi/6)}$ 的笛卡儿形式是什么? 根据欧拉恒等式, 为 $2(\cos(5\pi/6) + \mathrm{i}\sin(5\pi/6))$. 明白为什么要知道三角函数形式吗? 希望你能算出 $\cos(5\pi/6) = -\sqrt{3}/2$ 和 $\sin(5\pi/6) = 1/2$, 所以

$$2\mathrm{e}^{\mathrm{i}(5\pi/6)} = 2\left(\cos\left(\frac{5\pi}{6}\right) + \mathrm{i}\sin\left(\frac{5\pi}{6}\right)\right) = 2\left(-\frac{\sqrt{3}}{2} + \mathrm{i}\frac{1}{2}\right) = -\sqrt{3} + \mathrm{i}.$$

另一方面, 如 27.2.1 节所述, 由笛卡儿形式转换到极坐标形式要稍难一些. 在那节,

$$r = \sqrt{x^2 + y^2} \quad \text{和} \quad \tan(\theta) = \frac{y}{x},$$

其中舍去了可能的解 $r = -\sqrt{x^2 + y^2}$, 因为我们需要复数的 $r \geqslant 0$. 顺便说一下, 我们定义 z 的模为 $|z| = \sqrt{x^2 + y^2}$, 所以 r 等于 $|z|$. 因此模 $|z|$ 是从原点到点 z 的距离 (在复平面中). 角 θ 被称为 z 的辐角, 写为 $\arg(z)$. (通常要求 $0 \leqslant \arg(z) < 2\pi$ 以避免产生歧义.[①])

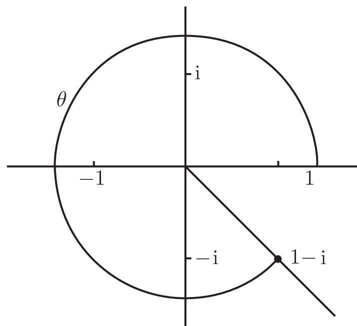


图 28-3

将 z 由笛卡儿坐标转换为极坐标, 只需用上面的公式求出 z 的模和辐角. (其实, 有时 z 的极坐标形式也被称为模–辐角式.) 例如, 如何将 $z = 1 - \mathrm{i}$ 转换成极坐标形式? 将 z 写作 $1 + (-1)\mathrm{i}$, 则需令上面公式中的 $x = 1$, $y = -1$. 事实上, 若 $z = r\mathrm{e}^{\mathrm{i}\theta}$, 则 $r = \sqrt{1^2 + (-1)^2} = \sqrt{2}$, $\tan(\theta) = (-1)/1 = -1$. 现在需要确定 θ 的正确值所在的象限. 最好的方法是画图, 如图 28-3 所示.

显然, 点 $(1, -1)$ 在第四象限, 所以 θ 一定等于 $7\pi/4$ (即 $\theta = -\pi/4$). 所以, 我们只需将 $r = \sqrt{2}$ 和 $\theta = 7\pi/4$ 一起代入 $r\mathrm{e}^{\mathrm{i}\theta}$ 而得到 $1 - \mathrm{i} = \sqrt{2}\mathrm{e}^{\mathrm{i}(7\pi/4)}$. (若用 $\theta = -\pi/4$, 将得到 $1 - \mathrm{i} = \sqrt{2}\mathrm{e}^{-\mathrm{i}(\pi/4)}$. 要知道在 θ 上加 2π 的任意整数倍都是正确的.)

① 这个条件也经常被写为 $-\pi < \arg(z) \leqslant \pi$.

让我们来看一对看似易混的例子. 首先, 如何写 2i 的极坐标形式? 考虑 2i 为 0+2i, 故它可由复平面上的点 $(0,2)$ 表示. 因此, 若 $2\mathrm{i} = r\mathrm{e}^{\mathrm{i}\theta}$, 则有 $r = \sqrt{0^2 + 2^2} = 2$, 而 $\tan(\theta) = 2/0$. 等一下, 不对, 0 不能作除数. 我们画个图来看看 θ 应该是什么 (如图 28-4 所示).

由图可知 $\theta = \dfrac{\pi}{2}$, 这与前面奇怪的 $\tan(\theta)$ 值一致, 因为 $\tan\left(\dfrac{\pi}{2}\right)$ 无定义. 因此, 我们有 $2\mathrm{i} = 2\mathrm{e}^{\mathrm{i}\pi/2}$. 当然, 这正是前一节的公式 $\mathrm{i} = \mathrm{e}^{\mathrm{i}\pi/2}$ 的 2 倍.

那将 -6 转换成极坐标形式呢? 现在我们将 -6 写为 $-6 + 0\mathrm{i}$, 可知

$$r = \sqrt{(-6)^2 + 0^2} = 6,\ \tan(\theta) = 0/(-6) = 0.$$

这意味着 θ 是 π 的整数倍, 为了确定它, 我们画另一幅图 (见图 28-5).

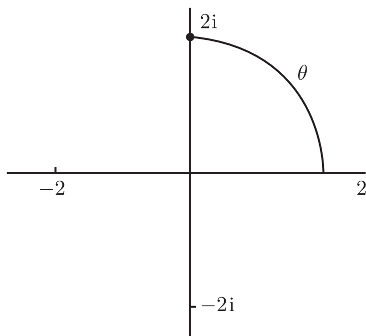


图 28-4

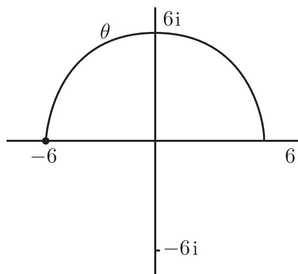


图 28-5

现在可知 $\theta = \pi$(或随你喜欢, 选 $-\pi$、3π, 甚至任何 π 的奇数倍). 因此, 我们有 $-6 = 6\mathrm{e}^{\mathrm{i}\pi}$. 顺便提一句, 若除以 6, 将得到令人惊异的公式 $\mathrm{e}^{\mathrm{i}\pi} = -1$, 该公式在上一节讨论过.

28.3 复数的高次幂

你究竟为什么要用极坐标形式呢? 一个原因是, 极坐标形式比较容易进行乘法和取幂运算. 设想用 $2\mathrm{e}^{-\mathrm{i}(3\pi/8)}$ 乘 $3\mathrm{e}^{\mathrm{i}\pi/4}$. 这个很简单, 只需用一般指数法则 (见 9.1.1 节) 得

$$(3\mathrm{e}^{\mathrm{i}\pi/4})(2\mathrm{e}^{-\mathrm{i}(3\pi/8)}) = 6\mathrm{e}^{\mathrm{i}(\pi/4-3\pi/8)} = 6\mathrm{e}^{-\mathrm{i}\pi/8}.$$

甚至更好的方式, 如取 $3\mathrm{e}^{\mathrm{i}\pi/4}$ 的 200 次幂, 就是

$$(3\mathrm{e}^{\mathrm{i}\pi/4})^{200} = 3^{200}\mathrm{e}^{\mathrm{i}(\pi/4)\times 200} = 3^{200}\mathrm{e}^{\mathrm{i}(50\pi)}.$$

事实上, 由欧拉恒等式可知, $\mathrm{e}^{\mathrm{i}(50\pi)} = \cos(50\pi) + \mathrm{i}\sin(50\pi)$. 由于 50π 是 2π 的整数倍, 所以有 $\cos(50\pi) = 1$ 和 $\sin(50\pi) = 0$, 故证得 $(3\mathrm{e}^{\mathrm{i}\pi/4})^{200} = 3^{200}$.

很多时候, 你想要的最终结果可能是笛卡儿形式的. 例如计算 $(1-\mathrm{i})^{99}$, 并给

出笛卡儿形式的结果. 将该式展开是很荒唐的, 所以我们不会那么做. 正确的方法是将 $1-\mathrm{i}$ 转换为极坐标形式, 取 99 次幂, 然后再转换回笛卡儿形式. 来看我们在前一节见到的极坐标形式的 $1-\mathrm{i}=\sqrt{2}\mathrm{e}^{\mathrm{i}(7\pi/4)}$, 所以有

$$(1-\mathrm{i})^{99}=(\sqrt{2}\mathrm{e}^{\mathrm{i}(7\pi/4)})^{99}=(2^{1/2})^{99}(\mathrm{e}^{\mathrm{i}(7\pi/4)})^{99}=2^{99/2}\mathrm{e}^{\mathrm{i}(693\pi/4)}.$$

现在要回到笛卡儿形式. 在转换之前, 我们来看 $\mathrm{e}^{\mathrm{i}(693\pi/4)}$. 这个分数 $693\pi/4$ 有点讨厌. 要知道 $\mathrm{e}^{\mathrm{i}\theta}$ 以关于 θ 的 2π 为周期, 所以把分数 $693\pi/4$ 的所有 2π 倍数去掉而不影响结果, 故有 $693/4=173\frac{1}{4}$, 小于这个数的最大偶数为 172, 且这两个数之差为 $173\frac{1}{4}-172=5/4$. 所以我们可将 $693\pi/4$ 看作 $172\pi+5\pi/4$. 因为 172π 是 2π 的整数倍 (这就是我们想要偶数的原因, 172 就是这种情况), 所以 $\mathrm{e}^{\mathrm{i}(693\pi/4)}=\mathrm{e}^{\mathrm{i}(5\pi/4)}$. 这样就好多了, 现在可将整个式子写成笛卡儿形式:

$$\begin{aligned}(1-\mathrm{i})^{99}&=2^{99/2}\mathrm{e}^{\mathrm{i}(693\pi/4)}=2^{99/2}\mathrm{e}^{\mathrm{i}(5\pi/4)}=2^{99/2}\left(\cos\left(\frac{5\pi}{4}\right)+\mathrm{i}\sin\left(\frac{5\pi}{4}\right)\right)\\&=2^{99/2}\left(-\frac{1}{\sqrt{2}}-\mathrm{i}\frac{1}{\sqrt{2}}\right).\end{aligned}$$

其实, 这个式子还可以进一步将 $1/\sqrt{2}$ 简写为 $2^{-1/2}$, 最终的结果为 $-2^{49}(1+\mathrm{i})$. 作为练习, 你可以对另一种极坐标形式 $1-\mathrm{i}=\sqrt{2}\mathrm{e}^{-\mathrm{i}\pi/4}$, 验证有相同的结果.

总之, 若取复数的高次幂, 首先将它转化为极坐标形式, 然后取幂. 求小于 θ 的 π 的最大偶数倍, 然后从 θ 中减去这个偶数倍, 且用所得新数来代换 θ. 最后, 换回笛卡儿形式.

28.4 解 $z^n = w$

我们来看一个复杂的主题: 如何解形如 $z^n=w$ 的方程, 其中 n 是整数, w 为复数. 这意味着要取 w 的 n 次方根, 但并不是简单的 $z=\sqrt[n]{w}$, 因为它没有告诉我们太多信息. 相反, 我们将直接求解. 因为极坐标形式的幂次很好算, 而它就是我们要用的.

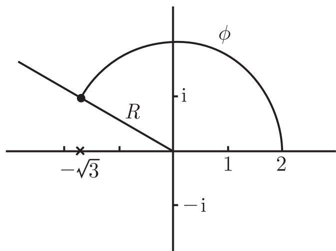


图 28-6

例如, 求解 $z^5=-\sqrt{3}+\mathrm{i}$, 我们应该同时用 z 和 $w=-\sqrt{3}+\mathrm{i}$ 的极坐标. 因为不知道 z 的值, 所以令 $z=r\mathrm{e}^{\mathrm{i}\theta}$. 要求 z, 只需求 r 和 θ. 对于 w, 我们写出 $-\sqrt{3}+\mathrm{i}=R\mathrm{e}^{\mathrm{i}\phi}$ 并求 R 和 ϕ. (这里用 R 和 ϕ 而不是 r 和 θ, 是因为后两个变量已经用于 z 了.) 现在, 我们来画出该情形对应的图 (见图 28-6).

所以我们有 $R=\sqrt{(-\sqrt{3})^2+1^2}=2$ 和 $\tan(\phi)=-1/\sqrt{3}$. 由于点在第二象限, ϕ 必为 $5\pi/6$. 太好了, 我们知道了极坐标形式下 $-\sqrt{3}+\mathrm{i}=2\mathrm{e}^{\mathrm{i}(5\pi/6)}$.

现在我们把注意力转移到方程 $z^5 = -\sqrt{3} + \mathrm{i}$ 上, 并将它们转换为极坐标形式. 在左边, 我们用 $r\mathrm{e}^{\mathrm{i}\theta}$ 代换 z, 得到 $z^5 = (r\mathrm{e}^{\mathrm{i}\theta})^5 = r^5\mathrm{e}^{\mathrm{i}(5\theta)}$; 而我们已经知道右边为 $2\mathrm{e}^{\mathrm{i}(5\pi/6)}$. 故方程变为

$$r^5\mathrm{e}^{\mathrm{i}(5\theta)} = 2\mathrm{e}^{\mathrm{i}(5\pi/6)}.$$

若两边同时取模, 可得 $r^5 = 2$ (因为若 A 是实数, $\mathrm{e}^{\mathrm{i}A}$ 的模总为 1). 然后我们可以消去 r^5 和 2, 因为它们相等, 从而得到 $\mathrm{e}^{\mathrm{i}(5\theta)} = \mathrm{e}^{\mathrm{i}(5\pi/6)}$. 我们已将上述方程分成了两个独立的方程:

$$r^5 = 2 \text{ 和 } \mathrm{e}^{\mathrm{i}(5\theta)} = \mathrm{e}^{\mathrm{i}(5\pi/6)}$$

第一个方程很容易求解: 只需取 5 次方根, 得到 $r = 2^{1/5}$, 这是合理的, 因为 r 是一个非负实数. 对于第二个方程, 你可能想说 $5\theta = 5\pi/6$, 但其实没那么简单. 记住, $\mathrm{e}^{\mathrm{i}\theta}$ 关于变量 θ 以 2π 为周期! 你可以通过下面这个重要原理来阐释这个结果, 这个原理你一定要着重记忆:

> 若 $\mathrm{e}^{\mathrm{i}A} = \mathrm{e}^{\mathrm{i}B}$ 对任意实数 A 和 B 成立, 则 $A = B + 2\pi k$, 其中 k 是整数.

该原理使我们转危为安. 由于 $\mathrm{e}^{\mathrm{i}(5\theta)} = \mathrm{e}^{\mathrm{i}(5\pi/6)}$, 因而我们运用该原理有

$$5\theta = \frac{5\pi}{6} + 2\pi k,$$

其中 k 为整数. 除以 5, 有

$$\theta = \frac{\pi}{6} + \frac{2\pi k}{5}.$$

看起来好像有无穷多个 θ 值, 因此方程有无穷多个 z 值. 然而, 外表是有欺骗性的! 你看, 由于 $n = 5$, 所以只需要用到 k 的前 5 个值, 即 $k = 0, 1, 2, 3, 4$. 我们一会儿会讨论其原因. 现在, 我们可以计算出 k 从 0 取到 4 时, θ 的值分别为

$$\frac{\pi}{6}, \qquad \left(\frac{\pi}{6} + \frac{2\pi}{5}\right) = \frac{17\pi}{30}, \qquad \left(\frac{\pi}{6} + \frac{4\pi}{5}\right) = \frac{29\pi}{30},$$
$$\left(\frac{\pi}{6} + \frac{6\pi}{5}\right) = \frac{41\pi}{30}, \qquad \left(\frac{\pi}{6} + \frac{8\pi}{5}\right) = \frac{53\pi}{30},$$

将 θ 的这些值和 $r = 2^{1/5}$ 代入方程 $z = r\mathrm{e}^{\mathrm{i}\theta}$, 可得

$$z = 2^{1/5}\mathrm{e}^{\mathrm{i}\pi/6}, \quad 2^{1/5}\mathrm{e}^{\mathrm{i}(17\pi/30)}, \quad 2^{1/5}\mathrm{e}^{\mathrm{i}(29\pi/30)}, \quad 2^{1/5}\mathrm{e}^{\mathrm{i}(41\pi/30)}, \quad 2^{1/5}\mathrm{e}^{\mathrm{i}(53\pi/30)}.$$

当然, 将这些解转换为笛卡儿形式就好了. 第一个解很容易:

$$2^{1/5}\mathrm{e}^{\mathrm{i}\pi/6} = 2^{1/5}\left(\cos\left(\frac{\pi}{6}\right) + \mathrm{i}\sin\left(\frac{\pi}{6}\right)\right) = 2^{1/5}\left(\frac{\sqrt{3}}{2} + \mathrm{i}\frac{1}{2}\right) = 2^{-4/5}(\sqrt{3} + \mathrm{i}).$$

其他的解看起来就没那么简单了. 例如, 第二个解

$$2^{1/5}\mathrm{e}^{\mathrm{i}(17\pi/30)} = 2^{1/5}\left(\cos\left(\frac{17\pi}{30}\right) + \mathrm{i}\sin\left(\frac{17\pi}{30}\right)\right)$$



并不容易化简. (你知道 $\cos(17\pi/30)$ 是多少吗? 我不知道, 但没必要算出来.) 我把用 (未化简的) 笛卡儿形式写出其他 3 个解的工作留给你完成.

现在, 我们来看为什么 k 只需从 0 取到 4, 而舍掉其他所有的 k 值. 我们来看当 $k=5$ 时会怎样. 运用前面的方程

$$\theta = \frac{\pi}{6} + \frac{2\pi k}{5},$$

当 $k=5$ 时, 我们有

$$\theta = \frac{\pi}{6} + \frac{2\pi \times 5}{5} = \frac{\pi}{6} + 2\pi.$$

这个结果当然与我们给出的其他 θ 值不一样, 但并没有出现不同的 z 值. 为什么? 因为

$$2^{1/5}e^{i(\pi/6+2\pi)} = 2^{1/5}e^{i(\pi/6)}.$$

即, 我们得到了一个与 $k=0$ 情形一样的解. 类似地, 若 $k=6$, 你应该得到与 $k=1$ 时一样的 z 值. 一般地, 每次将 k 加 5, 你就会再一次得到相同的 z 值. 所以, $k=0,5,10,\cdots$ 和 $k=-5,-10,-15,\cdots$ 一样有相同的解, 即 $z=2^{1/5}e^{i(\pi/6)}$. 类似地, $k=1,6,11,\cdots$ 和 $k=-4,-9,-14,\cdots$ 给出了相同的解. 其他 3 个解也一样. 你应该重视这个结果, 它在实践中很容易应用: 除非 $w=0$, 否则在 $k=0,1,\cdots,n-1$ 时, 方程 $z^n=w$ 有 n 个不同的解. 那些就是你要用到的 k 值. 我们的例子中 $n=5$, 所以只需要 $k=0,1,2,3,4$.

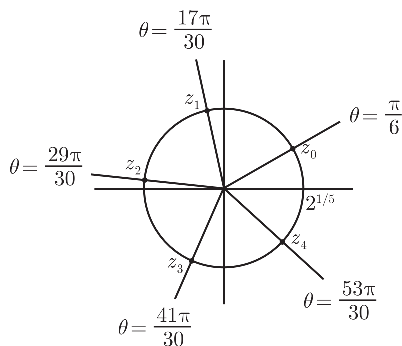


图 28-7

在复平面中描画所有的解会很有意思. 它们的模都为 $2^{1/5}$, 这意味着它们都在中心在原点、半径为 $2^{1/5}$ 单位的圆上. 而且, 连续解的辐角差 (即 θ 值) 为 $2\pi/5$, 它是整个圆周的五分之一. 这意味着所有的解均匀的分布在圆周上; 也就是说, 它们形成了一个规则的五边形 (解用 z_0 到 z_4 标记), 如图 28-7 所示.

一般地, 方程 $z^n=w$ 有 n 个解, 当画出这些解时, 它们的顶点形成了一个正 n 多边形. ($w=0$ 时例外, 这种情况下 $z=0$ 是唯一的解, 但它是 n 重的).



我们来给出求解 $z^n=w$ 的主要步骤:

(1) 将 $z=re^{i\theta}$ 写成极坐标. 则 $z^n=r^ne^{in\theta}$;

(2) 将 w 转化成极坐标. 我们设 $w=Re^{i\phi}$;

(3) 由于 $z^n=w$, 因而原方程可以写成 $r^ne^{in\theta}=Re^{i\phi}$, 这里的 n、R 和 ϕ 的值已知, 而 r 和 θ 是我们想求的 (所以它们作为变量出现);

(4) 分成两个方程: $r^n=R$ 和 $e^{in\theta}=e^{i\phi}$;

(5) 第一个方程很容易求解：取 n 次方根可得 $r = R^{1/n}$,

(6) 求解第二个方程可运用前面方框中的原理, 于是得 $n\theta = \phi + 2\pi k$, 其中 k 是整数;

(7) 用 n 除这个结果, 然后写出当 $k = 0, 1, 2, \cdots, n-1$ 时的所有不同的 θ 值;

(8) 将 r 值和不同的 θ 值代入 $z = re^{i\theta}$, 得到 n 个不同的 z 值, 即为解;

(9) 若有必要, 将每个解转换成笛卡儿坐标形式.

我们再看一个例子：i 的三次方根是多少？ 这个问题需求解方程 $z^3 = i$. 我们先写出 $z = re^{i\theta}$, 则 $z^3 = r^3e^{i(3\theta)}$ (第 1 步). 现在, 我们需要将 i 转换成极坐标 (第 2 步), 我们已经知道 $i = e^{i\pi/2}$. 由于 $z^3 = i$, 因而我们有 $r^3e^{i(3\theta)} = 1e^{i\pi/2}$ (第 3 步). 这就导出方程 $r^3 = 1$ 和 $e^{i(3\theta)} = e^{i\pi/2}$ (第 4 步). 对第一个方程取三次方根, 可得 $r = 1$ (第 5 步), 再根据重要原理, 第二个方程解出 $3\theta = \pi/2 + 2\pi k$, 其中 k 是整数 (第 6 步). 这与 $\theta = \pi/6 + 2\pi k/3$ 等价, 由于这个问题中的 $n = 3$, 所以我们只需取 $k = 0, 1, 2$, 于是得到

$$\theta = \frac{\pi}{6}, \qquad \left(\frac{\pi}{6} + \frac{2\pi}{3}\right) = \frac{5\pi}{6}, \qquad \left(\frac{\pi}{6} + \frac{4\pi}{3}\right) = \frac{3\pi}{2}$$

(第 7 步). 继而推导出 z 的三个可能值为

$$z = e^{i\pi/6}, \quad e^{i(5\pi/6)}, \quad e^{i(3\pi/2)}$$

(第 8 步). 最后, 我们应该将这些解转换为笛卡儿形式 (第 9 步). 第一个解为

$$z = e^{i\pi/6} = \cos\left(\frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{6}\right) = \frac{\sqrt{3}}{2} + i\frac{1}{2}.$$

第二个解为

$$z = e^{i(5\pi/6)} = \cos\left(\frac{5\pi}{6}\right) + i\sin\left(\frac{5\pi}{6}\right) = -\frac{\sqrt{3}}{2} + i\frac{1}{2}.$$

第三个解为

$$z = e^{i(3\pi/2)} = \cos\left(\frac{3\pi}{2}\right) + i\sin\left(\frac{3\pi}{2}\right) = 0 - i(1) = -i.$$

我们画出这三个解, 从图 28-8 中可以看出, 它们的确形成了一个等边三角形.

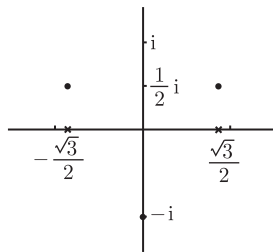


图 28-8

一些变式

假设要求解方程 $(z-2)^3 = i$. 没问题, 只需令 $Z = z - 2$, 则方程为 $Z^3 = i$. 像上节我们所做的那样来解这个方程, 可得

$$Z = z - 2 = \frac{\sqrt{3}}{2} + i\frac{1}{2}, \qquad -\frac{\sqrt{3}}{2} + i\frac{1}{2}, \qquad -i.$$

最后, 两边同时加 2, 可得

$$z = 2 + \frac{\sqrt{3}}{2} + i\frac{1}{2}, \qquad 2 - \frac{\sqrt{3}}{2} + i\frac{1}{2}, \qquad 2 - i.$$

一点都不难. 不过这里有个较难的, 我们来试着求解二次方程

$$z^2 + \frac{1}{\sqrt{2}}z - \frac{\sqrt{3}\mathrm{i}}{8} = 0.$$

运用二次公式, 我们得到

$$z = \frac{\frac{-1}{\sqrt{2}} \pm \sqrt{\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}}}{2}.$$

这个结果是正确的, 但它不是笛卡儿形式 (也不是极坐标形式), 所以我们要试着化简它. 我们需要求复数 $\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$ 的二次方根. 怎么做呢? 求解方程 $Z^2 = \frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$. 根据前面的步骤, 我们写出 $Z = r\mathrm{e}^{\mathrm{i}\theta}$, 极坐标形式为 $\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2} = \mathrm{e}^{\mathrm{i}\pi/3}$, 可自行证明. 所以我们的方程变为 $r^2\mathrm{e}^{\mathrm{i}2\theta} = \mathrm{e}^{\mathrm{i}\pi/3}$. 这意味着 $r^2 = 1$ 和 $2\theta = \pi/3 + 2\pi k$, 其中 $k = 0$ 或 1. (要记住重要原理!) 所以, 我们有 $r = 1$ 和 $\theta = \pi/6$ 或 $7\pi/6$, 这意味着 $Z = \mathrm{e}^{\mathrm{i}\pi/6}$ 或 $Z = \mathrm{e}^{\mathrm{i}7\pi/6}$. 再一次, 需要验证, 它们对应的笛卡儿形式分别为 $Z = \frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}$ 或 $Z = -\frac{\sqrt{3}}{2} - \frac{1}{2}\mathrm{i}$. 最后, 我们可以在上面 z 的方程中用 $\pm\left(\frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}\right)$ 来代换 $\pm\sqrt{\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}}$, 得到

$$z = \frac{-\frac{1}{\sqrt{2}} \pm \left(\frac{\sqrt{3}}{2} + \frac{1}{2}\mathrm{i}\right)}{2},$$

化简为

$$z = -\frac{1}{2\sqrt{2}} + \frac{\sqrt{3}}{4} + \frac{\mathrm{i}}{4} \quad \text{或} \quad -\frac{1}{2\sqrt{2}} - \frac{\sqrt{3}}{4} - \frac{\mathrm{i}}{4}$$

来看另一个例子. 如何在复数域上将 $(z^4 - z^2 + 1)$ 因式分解? 在实数域上又如何呢? 在第一个情形中, 我们只需求出方程 $z^4 - z^2 + 1 = 0$ 的所有复数解, 共 4 个. 为了求解, 我们首先需要知道这个方程其实是 z^2 的二次方程. 我们令 $Z = z^2$, 则方程变为 $Z^2 - Z + 1 = 0$. 运用二次公式求解得

$$Z = z^2 = \frac{1 \pm \sqrt{-3}}{2} = \frac{1}{2} \pm \mathrm{i}\frac{\sqrt{3}}{2}.$$

我们需求 $\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$ 和 $\frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}$ 的二次方根. 我们在前个例子中已经完成了对第一个数的求解, 你可以按相同的步骤来处理第二个数, 足够简单. 这两个数各有两个平方根, 算出来为

$$\frac{\sqrt{3} + \mathrm{i}}{2}, \quad \frac{-\sqrt{3} - \mathrm{i}}{2}, \quad \frac{-\sqrt{3} + \mathrm{i}}{2}, \quad \frac{\sqrt{3} - \mathrm{i}}{2}.$$

这些是 $z^4 - z^2 + 1 = 0$ 的解. 由此, 我们可将 $z^4 - z^2 + 1$ 因式分解为

$$z^4-z^2+1=\left(z-\frac{\sqrt{3}+\mathrm{i}}{2}\right)\left(z-\frac{\sqrt{3}-\mathrm{i}}{2}\right)\left(z-\frac{-\sqrt{3}+\mathrm{i}}{2}\right)\left(z-\frac{-\sqrt{3}-\mathrm{i}}{2}\right).$$

这是复因式分解. 为了求出实因式分解, 我们需要运用一个事实：若 w 为任意复数, 则 $(z-w)(z-\bar{w})$ 相乘有实系数. 事实上, 你会得到 $z^2-(w+\bar{w})z+w\bar{w}$, 易知 $w+\bar{w}=2\mathrm{Re}(w)$(为实数), 而我们已知 $w\bar{w}=|w|^2$, 也是实数. 不知道你是否注意到, 我已将四个因式巧妙地分了组, 使得将前两个因式相乘时得到

$$\begin{aligned}\left(z-\frac{\sqrt{3}+\mathrm{i}}{2}\right)\left(z-\frac{\sqrt{3}-\mathrm{i}}{2}\right)&=z^2-\left(\frac{\sqrt{3}+\mathrm{i}}{2}+\frac{\sqrt{3}-\mathrm{i}}{2}\right)z+\left(\frac{\sqrt{3}+\mathrm{i}}{2}\right)\left(\frac{\sqrt{3}-\mathrm{i}}{2}\right)\\&=z^2-\sqrt{3}z+1.\end{aligned}$$

类似地, 你可以验证后两个因式相乘得到的是 $z^2+\sqrt{3}z+1$. 故结论是

$$z^4-z^2+1=(z^2-\sqrt{3}z+1)(z^2+\sqrt{3}z+1).$$

注意, 这里没有任何复数. 然而, 这个例子若不用因式来计算, 会相当棘手.

28.5　解 $\mathrm{e}^z=w$

现在该讨论如何对给定的 w 求解形如 $\mathrm{e}^z=w$ 的方程了. 要是能写成 $z=\ln(w)$ 就好了, 但帮助并不大. 例如, $\ln(-\sqrt{3}+\mathrm{i})$ 是多少呢？让我们来回答这个问题.

幸运的是, 求解 $\mathrm{e}^z=w$ 并不比求解 $z^n=w$ 难多少, 事实上, 若说有什么区别的话, 就是求解更简单. 在讨论解法之前, 我们需要更多地理解 e^z. 我们来看如果写成 $z=x+\mathrm{i}y$ 会发生什么. 我们得到

$$\mathrm{e}^z=\mathrm{e}^{x+\mathrm{i}y}=\mathrm{e}^x\mathrm{e}^{\mathrm{i}y}.$$

那又怎样？这里的关键是, 这已经是极坐标形式了. 模为 e^x, 辐角为 y. 如果你愿意, 也可写为 $r=\mathrm{e}^x$ (记住, e^x 是正实数), $\theta=y$. 这意味着, 若 z 的笛卡儿形式是 $x+\mathrm{i}y$, 则 e^z **自动**有极坐标形式 $\mathrm{e}^z=\mathrm{e}^x\mathrm{e}^{\mathrm{i}y}$. 所以, 求解 $\mathrm{e}^z=w$ 和 $z^n=w$ 时的主要区别是, 前个问题中不必将 z 写成极坐标形式, 而后个问题中需要这样做. 这个问题还有个副产物, 即方程 $\mathrm{e}^z=w$ 有无穷多个解 (除非 $w=0$, 在这种情况下方程无解).



我们来求解 $\mathrm{e}^z=-\sqrt{3}+\mathrm{i}$. 我们已经将右边转换成了极坐标 $2\mathrm{e}^{\mathrm{i}(5\pi/6)}$ (参见 28.4 节). 为了处理左边, 要将 $z=x+\mathrm{i}y$ 写成笛卡儿坐标, 所以 $\mathrm{e}^z=\mathrm{e}^x\mathrm{e}^{\mathrm{i}y}$. 因此, 将原方程转换成极坐标形式, 可得

$$\mathrm{e}^x\mathrm{e}^{\mathrm{i}y}=2\mathrm{e}^{\mathrm{i}(5\pi/6)}.$$

现在分成两个方程;

$$\mathrm{e}^x=2 \text{ 和 } \mathrm{e}^{\mathrm{i}y}=\mathrm{e}^{\mathrm{i}(5\pi/6)}.$$

为求解第一个方程, 我们需要取对数, 可知 $x = \ln(2)$. 对第二个方程, 我们运用重要原理可得 $y = 5\pi/6 + 2\pi k$, 其中 k 是整数. 最后, 将这些值代入 $z = x + iy$, 可得

$$z = \ln(2) + i\left(\frac{5\pi}{6} + 2\pi k\right),$$

其中 k 是任意整数. 在本例中, 不同的 k 值有不同的 z 值, 所以我们需要用到所有的 k 值. 我们来画出对应 $k = -2, -1, 0, 1, 2$ 的 z 值 (为了清晰起见, 我们对两坐标轴运用了不同的尺度), 如图 28-9 所示.

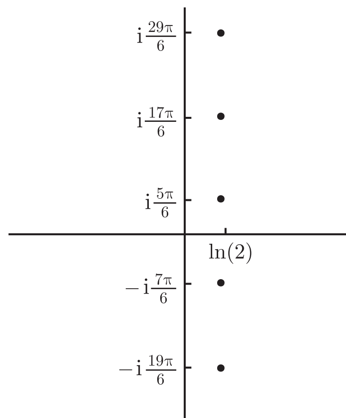


图 28-9

所以, 解都在垂直线 $x = \ln(2)$ 上均匀分布. 顺便提一句, 这意味着它们形成复数的等差数列. 虽然图中只显示了 5 个解, 但你需要记住, 方程 $e^z = -\sqrt{3} + i$ 其实有无穷多个解.

我们再来看个例子. 假设要求解 $e^{2iz+3} = i$. 指数 $2iz + 3$ 让这个例显得比前一例复杂一点, 但其实没那么糟. 我们已经知道右边的极坐标是 $e^{i\pi/2}$, 那左边呢? 同样, 我们写出 $z = x + iy$, 不过现在需令 $2iz + 3 = 2i(x + iy) + 3 = (-2y + 3) + i(2x)$. 所以, 左边的极坐标形式为

$$e^{2iz+3} = e^{-2y+3}e^{i(2x)}.$$

注意 i 的因式是如何改变实部和虚部的 (还有 y 的符号). 不管怎样, 将方程 $e^{2iz+3} = i$ 转换成极坐标形式, 我们有

$$e^{-2y+3}e^{i(2x)} = 1e^{i\pi/2}.$$

由此可推出方程

$$e^{-2y+3} = 1 \text{ 和 } e^{i(2x)} = e^{i\pi/2}.$$

为解第一个方程, 取对数可得 $-2y + 3 = \ln(1) = 0$, 所以 $y = \dfrac{3}{2}$. 为解第二个方程, 运用重要原理可得 $2x = \pi/2 + 2\pi k$, 其中 k 是整数. 这意味着 $x = \pi/4 + \pi k$, 所以由 $z = x + iy$, 我们有

$$z = \frac{\pi}{4} + \pi k + \frac{3}{2}i,$$

其中 k 是整数. 我们画出 $k = -2, -1, 0, 1, 2$ 的解来看看它们是什么样的, 如图 28-10 所示. 记住, 这些解只是无穷多个解中的 5 个. 同样, 这些解为等差数列, 不过这次分布在水平线 $y = \dfrac{3}{2}$ 上.

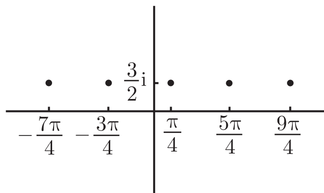


图 28-10

28.6　一些三角级数

三角级数是系数为 $\{a_n\}$ 和 $\{b_n\}$, 形如

$$\sum_{n=0}^{\infty}(a_n\cos(n\theta)+b_n\sin(n\theta))$$

的级数. 在本节, 我们将会看到这类可被化简的级数.

例如, 考虑三角级数

$$\sum_{n=0}^{\infty}\frac{\sin(n\theta)}{n!},$$

其中 θ 为实数. 注意, 这不是一个关于 θ 的幂级数, 因为 $\sin(n\theta)$ 不是 θ 的幂. 另一方面, 我们可巧妙地运用互补级数

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{n!},$$

将整个级数转换成一个幂级数. 实际上, 我们可以立刻找到两个级数. 关键点就是欧拉恒等式. 要仔细看, 因为这个技巧很巧妙. 我们将它们组合成

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{n!}+\mathrm{i}\sum_{n=0}^{\infty}\frac{\sin(n\theta)}{n!},$$

就能快速找到这两个级数. 好吧, 这就是一个级数加上另一个级数乘以 i. 那又怎样？整理和式[①], 然后运用欧拉恒等式, 它可化简为

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)+\mathrm{i}\sin(n\theta)}{n!}=\sum_{n=0}^{\infty}\frac{\mathrm{e}^{\mathrm{i}n\theta}}{n!}.$$

最后, 运用指数规则将 $\mathrm{e}^{\mathrm{i}n\theta}$ 写成 $(\mathrm{e}^{\mathrm{i}\theta})^n$, 和式变为

$$\sum_{n=0}^{\infty}\frac{(\mathrm{e}^{\mathrm{i}\theta})^n}{n!}.$$

这个和式看起来很熟悉. 其实, 我们在 28.1.1 节见过, 对所有复数 z 有

$$\sum_{n=0}^{\infty}\frac{z^n}{n!}=\mathrm{e}^z.$$

现在我们只需代入 $z=\mathrm{e}^{\mathrm{i}\theta}$ 得到

$$\sum_{n=0}^{\infty}\frac{(\mathrm{e}^{\mathrm{i}\theta})^n}{n!}=\mathrm{e}^{\mathrm{e}^{\mathrm{i}\theta}}$$

如果你能明白上述推理, 就应该能明白我们已经证明了

① 这需要一些理由. 事实上, 一切都很顺利, 因为两个级数均绝对收敛.

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{n!}+\mathrm{i}\sum_{n=0}^{\infty}\frac{\sin(n\theta)}{n!}=\mathrm{e}^{\mathrm{e}^{\mathrm{i}\theta}}.$$

现在做什么？我们需要将右边转换成笛卡儿形式. 为此, 写出 $\mathrm{e}^{\mathrm{i}\theta}=\cos(\theta)+\mathrm{i}\sin(\theta)$, 故

$$\mathrm{e}^{\mathrm{e}^{\mathrm{i}\theta}}=\mathrm{e}^{\cos(\theta)+\mathrm{i}\sin(\theta)}=\mathrm{e}^{\cos(\theta)}\mathrm{e}^{\mathrm{i}\sin(\theta)}.$$

这是一个好的开端 —— 这是 $\mathrm{e}^{\mathrm{e}^{\mathrm{i}\theta}}$ 的极坐标形式. 为了得到笛卡儿形式, 我们需要将 $\mathrm{e}^{\mathrm{i}\sin(\theta)}$ 转换成 $\cos(\sin(\theta))+\mathrm{i}\sin(\sin(\theta))$. 综上, 我们可得

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{n!}+\mathrm{i}\sum_{n=0}^{\infty}\frac{\sin(n\theta)}{n!}=\mathrm{e}^{\cos(\theta)}\cos(\sin(\theta))+\mathrm{i}\mathrm{e}^{\cos(\theta)}\sin(\sin(\theta)).$$

现在, 若两个复数相等, 则它们的实部必须相等, 虚部也必须相等. 由此可推出下面的两个等式, 它们对所有实数 θ 都成立：

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{n!}=\mathrm{e}^{\cos(\theta)}\cos(\sin(\theta))\quad 和\quad \sum_{n=0}^{\infty}\frac{\sin(n\theta)}{n!}=\mathrm{e}^{\cos(\theta)}\sin(\sin(\theta)).$$



确实不容易, 但这些基本上都是你必须做的. 我再举个例子, 但这次不会给出任何解释. 你的任务是跟随每一步并给出相应的解释. 这个例子是求

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{3^n}\quad 和\quad \sum_{n=0}^{\infty}\frac{\sin(n\theta)}{3^n}.$$

遵照前面例子的求解过程, 我们有

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{3^n}+\mathrm{i}\sum_{n=0}^{\infty}\frac{\sin(n\theta)}{3^n}&=\sum_{n=0}^{\infty}\frac{\cos(n\theta)+\mathrm{i}\sin(n\theta)}{3^n}\\&=\sum_{n=0}^{\infty}\frac{\mathrm{e}^{\mathrm{i}n\theta}}{3^n}=\sum_{n=0}^{\infty}\frac{\mathrm{e}^{(\mathrm{i}\theta)^n}}{3^n}=\sum_{n\to 0}^{\infty}\left(\frac{\mathrm{e}^{\mathrm{i}\theta}}{3}\right)^n.\end{aligned}$$

这是一个公比为 $\mathrm{e}^{\mathrm{i}\theta}/3$ 的几何级数. 最后的数为极坐标形式, 模为 $1/3<1$, 所以该几何级数收敛. 根据几何级数求和公式 (参见 23.1 节), 我们有

$$\sum_{n=0}^{\infty}\left(\frac{\mathrm{e}^{\mathrm{i}\theta}}{3}\right)^n=\frac{1}{1-\dfrac{1}{3}\mathrm{e}^{\mathrm{i}\theta}}.$$



现在我们要将这个结果转换成笛卡儿坐标, 这个任务有点令人讨厌. 首先, 试一下看能否完成. 若不能, 至少应该试着理解下面的步骤：

$$\begin{aligned}
\frac{1}{1-\frac{1}{3}\mathrm{e}^{\mathrm{i}\theta}} &= \frac{1}{1-\frac{1}{3}\cos(\theta)-\mathrm{i}\frac{1}{3}\sin(\theta)} \\
&= \frac{1}{1-\frac{1}{3}\cos(\theta)-\mathrm{i}\frac{1}{3}\sin(\theta)}\cdot\frac{1-\frac{1}{3}\cos(\theta)+\mathrm{i}\frac{1}{3}\sin(\theta)}{1-\frac{1}{3}\cos(\theta)+\mathrm{i}\frac{1}{3}\sin(\theta)} \\
&= \frac{1-\frac{1}{3}\cos(\theta)+\mathrm{i}\frac{1}{3}\sin(\theta)}{\left(1-\frac{1}{3}\cos(\theta)\right)^2+\left(\frac{1}{3}\sin(\theta)\right)^2} \\
&= \frac{1-\frac{1}{3}\cos(\theta)+\mathrm{i}\frac{1}{3}\sin(\theta)}{1-\frac{2}{3}\cos(\theta)+\frac{1}{9}\cos^2(\theta)+\frac{1}{9}\sin^2(\theta)} \\
&= \frac{1-\frac{1}{3}\cos(\theta)+\mathrm{i}\frac{1}{3}\sin(\theta)}{1-\frac{2}{3}\cos(\theta)+\frac{1}{9}} \\
&= \frac{9-3\cos(\theta)+\mathrm{i}3\sin(\theta)}{10-6\cos(\theta)} \\
&= \frac{9-3\cos(\theta)}{10-6\cos(\theta)}+\mathrm{i}\frac{3\sin(\theta)}{10-6\cos(\theta)}.
\end{aligned}$$

这之后, 我们完全可以得出

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{3^n}+\mathrm{i}\sum_{n=0}^{\infty}\frac{\sin(n\theta)}{3^n}=\frac{9-3\cos(\theta)}{10-6\cos(\theta)}+\mathrm{i}\frac{3\sin(\theta)}{10-6\cos(\theta)}.$$

由于实部和虚部必须相等, 可推出对所有的实数 θ 有

$$\sum_{n=0}^{\infty}\frac{\cos(n\theta)}{3^n}=\frac{9-3\cos(\theta)}{10-6\cos(\theta)} \quad 和 \quad \sum_{n=0}^{\infty}\frac{\sin(n\theta)}{3^n}=\frac{3\sin(\theta)}{10-6\cos(\theta)}$$

如你所见, 这些问题相当难!

28.7 欧拉恒等式和幂级数

在这章最后, 我们来看看用幂级数证明欧拉恒等式

$$\mathrm{e}^{\mathrm{i}\theta}=\cos(\theta)+\mathrm{i}\sin(\theta).$$

根据 28.1.1 节对 e^z 的定义, 将 z 替换为 $\mathrm{i}\theta$, 可以得到

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}\theta} &= 1+(\mathrm{i}\theta)+\frac{(\mathrm{i}\theta)^2}{2!}+\frac{(\mathrm{i}\theta)^3}{3!}+\frac{(\mathrm{i}\theta)^4}{4!}+\frac{(\mathrm{i}\theta)^5}{5!}+\frac{(\mathrm{i}\theta)^6}{6!}+\frac{(\mathrm{i}\theta)^7}{7!}+\cdots \\
&= 1+\mathrm{i}\theta-\frac{\theta^2}{2!}-\mathrm{i}\frac{\theta^3}{3!}+\frac{\theta^4}{4!}+\mathrm{i}\frac{\theta^5}{5!}-\frac{\theta^6}{6!}-\mathrm{i}\frac{\theta^7}{7!}+\cdots
\end{aligned}$$

由于 i 的幂在值 $1, \mathrm{i}, -1, -\mathrm{i}$ 间持续循环, 因而可推导出上述级数的偶次幂都有实系数, 而奇次幂都有虚系数. 另外, 隔项偶次幂项为负, 其余为正; 奇次幂项同理. 所以 $\mathrm{e}^{\mathrm{i}\theta}$ 的实部为

$$1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots = \cos(\theta),$$

虚部为

$$\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots = \sin(\theta).$$

(回顾一下这些麦克劳林级数, 参见 26.2 节.) 由最后的等式, 可推出 $\mathrm{e}^{\mathrm{i}\theta} = \cos(\theta) + \mathrm{i}\sin(\theta)$.

第 29 章　体积、弧长和表面积

我们已经用定积分求过面积. 现在我们将用它们来求体积、弧长和表面积. 对于体积和表面积, 我们将特别关注平面区域绕某轴旋转一周得到的立体, 这类立体被称为*旋转体*. 对于体积, 我们会讨论一些更一般的立体. 这里是本章内容的计划:

- 圆盘法和壳法求体积;
- 求更一般立体的体积;
- 求光滑曲线的弧长和带参数的质点速率;
- 求旋转体的表面积.

29.1　旋转体的体积

我们从求旋转体体积开始. 平面上有某个区域, 也有某个轴, 立体由该区域关于轴旋转得到. 为便于研究, 我们假设这些轴总是平行于 x 轴或 y 轴. (也可能有斜轴, 不过这会很麻烦, 除非用线性代数里的方法.)

在我们带上 3D 眼镜之前, 先来回顾一下定积分的原理. 我们在第 16 章讲过该原理, 这里快速回顾一下其主要思想. 我们先看求曲线

$$y = \sqrt{1-(x-3)^2}$$

下方、x 轴上方的区域面积. 它看起来像什么? 如果我们将方程平方并重整将得到 $(x-3)^2+y^2=1$, 它的图像是以 $(3,0)$ 为圆心、1 为半径的圆, 所以这个函数是圆的上半部分, 如图 29-1 所示.

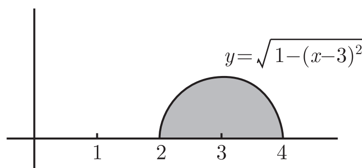


图　29-1

根据定积分的定义, 我们知道阴影区域的面积 (平方单位) 是

$$\int_2^4 \sqrt{1-(x-3)^2}\mathrm{d}x,$$

也可写作 $\int_2^4 y\mathrm{d}x$.

另一方面, 若使用黎曼和求该半圆的面积, 我们需将 x 轴上的底分割成小段, 然后将这些小段向上延伸为小条. 这些小条的宽度不必相同, 唯一需要确定的是每个小条的顶部要与曲线的某处相切 (即小条的某个角要触到曲线). 这些小条的面积和很容易求出, 因为它不过是矩形的面积和. 这个面积是半圆真正面积的近似, 小

条越细, 近似越好, 如图 29-2 所示.

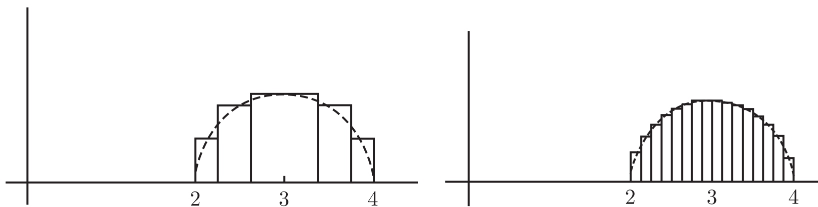


图 29-2

我们来看其中的一个小条. 为了便于讨论, 我们假设小条的左上角在曲线上. 如 16.4 节所述, 选择哪个小条都可以, 只要所有小条的顶部穿过曲线. 现在来看如图 29-3 所示的小条.

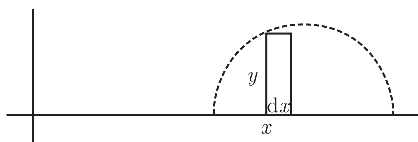


图 29-3

这个矩形小条的底宽 $\mathrm{d}x$ 个单位, 高 y 个单位, 它的面积是 $y\mathrm{d}x$ 平方单位. 现在我们要做的就是将所有小条的面积加起来, 同时令最大的底宽趋于 0. 积分符号的优势在于, 你只需将积分号写在小条面积的前面, 并给出正确的界. 在我们的例子中, x 在区间 $[2,4]$ 内, 一个小条的面积是 $y\mathrm{d}x$ 平方单位,所以**所有**小条的面积 —— 在小条最大的底宽趋于 0 时 —— 是 $\int_2^4 y\mathrm{d}x$ 平方单位.

所以, 其模式是这样的: 我们在 x 轴上的点 x 处取宽 $\mathrm{d}x$ 个单位、高 y 个单位的小条, 算出它的面积, 然后将积分号放在前面来得到要求的整个面积. 这种方法不仅适用于求面积, 也可用于求体积. 特别地, 让我们来看看它怎么通过两种不同的方法 —— 圆盘法和壳法, 求解旋转体的体积.

29.1.1 圆盘法

假设我们绕 x 轴旋转上节中的半圆, 这将会得到一个球. (能明白为什么吗?) 我们来尝试求体积. 我们从图 29-3 的小条开始, 然后这个小条关于 x 轴旋转得到图 29-4.

这是一个宽 $\mathrm{d}x$ 个单位、半径为 y 个单位的薄盘. 我们可把它看作一个圆柱体, 其半径为 y 个单位, 高为 $\mathrm{d}x$ 个单位. 由于半径为 r 单位、高为 h 个单位的圆柱体的体积为 $\pi r^2 h$ 立方单位, 所以小薄盘的体积是 $\pi y^2\mathrm{d}x$ 立方单位. 现在, 我们取一些小条, 它们的底是区间 $[2,4]$ 的一小段, 再让它们绕 x 轴旋转. 例如, 我们取 5 个小条并让其旋转, 就会得到图 29-5 所示的结果.

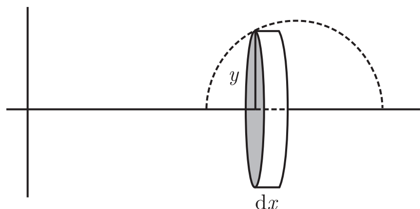
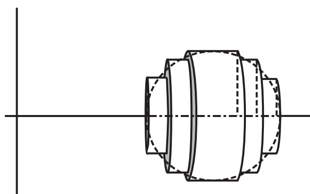


图　29-4


图　29-5

与完美的球相比, 图 29-5 的这个物体很蹩脚, 不过它的体积却是球体积的一个相当好的近似. 所用的圆盘越薄, 近似就越好. 在极限中, 当圆盘的最大厚度趋于 0 时, 近似变得完美：全部圆盘的总体积趋于球的体积. 同样地, 实现 “将所有圆盘体积加起来, 同时令圆盘最大厚度趋于 0” 的思想, 就是取任意一个圆盘的体积 ($\pi y^2\mathrm{d}x$ 立方单位) 并在我们需要的区间上进行积分. 在我们的例子中, $y=\sqrt{1-(x-3)^2}$ 且 x 从 2 取到 4, 所以我们有

$$V=\int_2^4 \pi y^2\mathrm{d}x=\pi\int_2^4(1-(x-3)^2)\mathrm{d}x$$

算出来的体积是 $\dfrac{4\pi}{3}$ 立方单位 (试一下!), 正是我们的预期, 因为我们讨论的是半径为 1 的球. 我们所用的方法称为*圆盘法*, 也称为*切片法*.

29.1.2　壳法

现在, 假设我们让图 29-1 的半圆形区域绕 y 轴旋转. 想想, 你会得到什么 —— 面包圈的上半部分 (没有罂粟籽). 我们再次用小细条来近似半圆, 但这次要关于 y 轴而不是 x 轴旋转每个小条. 如我们之前所见, 一个小条如图 29-6 所示.

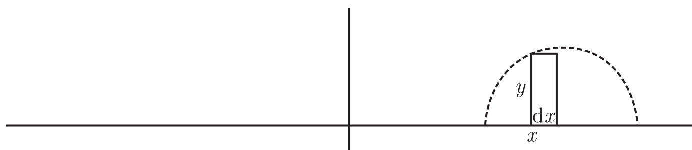


图　29-6

它关于 y 轴旋转, 得到的不是一个圆盘, 而是一个柱壳, 如图 29-7 所示.

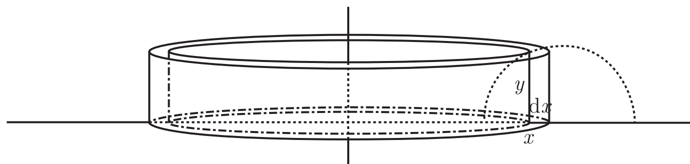


图　29-7

我们将用一系列的壳来近似那半个面包圈, 然后令壳的最大厚度递减趋于 0.

例如, 如果与前节一样, 用 5 个小条来近似区域, 那么会得到图 29-8 所示的东西.

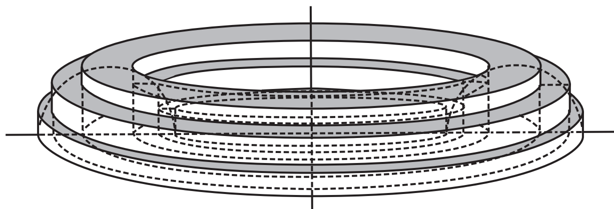


图 29-8

这个怪怪的立体真的是半个成块的面包圈, 但它的体积相当接近我们的所求. 壳的最大厚度越小, 近似就越好. 与前面一样, 积分关注于所有柱壳的体积之和及壳的最大厚度趋于 0 的极限.

首先, 我们要求一个普通壳的体积. 最简单的方法是把壳看作是一个没有底和顶的薄金属罐. 如图 29-7 所示, 罐子的高是 y 个单位, 半径是 x 个单位, 厚度是 $\mathrm{d}x$ 单位. 想象一下, 用锋利的剪刀沿罐子的一边剪开, 将其打开并平铺成一个薄薄的矩形状金属片. 当然它不是一个矩形. 要知道, 矩形是一个二维的物体, 而展开的罐子是三维的 —— 罐子虽然很薄, 但还是有厚度的. (甚至一片纸也有厚度, 否则大量的纸堆叠起来可能还是很薄很薄.) 它甚至不是一个长方体, 因为罐子的内半径不等于外半径. 但关键是, 它**几乎**是一个长方体. 罐子越薄, 就越接近于一个长方体, 当我们最后 (用积分) 求极限时,一切就都算出来了[①]. 所以, 理想的罐子展开图为图 29-9. 其厚度为 $\mathrm{d}x$ 单位, 剪开后高度仍为柱壳的高, 即 y 个单位. 那长边呢? 它等于壳的周长 (想一想), 即 $2\pi x$ 个单位, 因为壳的半径基本上是 x 个单位, 所以其体积与 $2\pi xy\mathrm{d}x$ 立方单位很接近. 现在我们所要做的, 就是从 $x=2$ 到 $x=4$ 进行积分来看半个面包圈的体积 (立方单位):

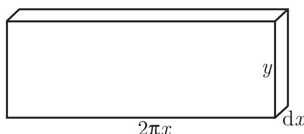


图 29-9

$$\int_2^4 2\pi xy\mathrm{d}x = 2\pi\int_2^4 x\sqrt{1-(x-3)^2}\mathrm{d}x.$$

太棒了! 我们已将问题简化为求定积分的值, 不过这个积分还是有点杂乱. 先做代换 $t=x-3$, 则 $\mathrm{d}t=\mathrm{d}x$. 同样, 当 $x=2$, 我们有 $t=-1$; 当 $x=4$, 我们可知 $t=1$. 所以, 用 t 表示, 积分变为

① 更一般的, 我们把壳的体积看成外面壳 (半径为 $x+\mathrm{d}x$ 个单位) 与里面壳 (半径为 x 个单位) 的体积差. 两个壳都有 y 单位高, 所以壳的体积为 $\pi y((x+\mathrm{d}x)^2-x^2)$, 化简为 $2\pi xy\mathrm{d}x+\pi y(\mathrm{d}x)^2$ 立方单位. 求积后, 第二项由于可忽略量 $(\mathrm{d}x)^2$ 而为 0.

$$2\pi\int_{-1}^{1}(t+3)\sqrt{1-t^2}\mathrm{d}t = 2\pi\left(\int_{-1}^{1}t\sqrt{1-t^2}\mathrm{d}t + 3\int_{-1}^{1}\sqrt{1-t^2}\mathrm{d}t\right).$$

第一个积分可通过做代换 $u=1-t^2$ 求解, 第二个积分可用三角换元求解. 有个更好的求解方法. 注意第一个积分其实为 0, 因为被积函数是关于 t 的奇函数, 且积分区域 $[-1,1]$ 关于 $t=0$ 对称. (我们在 18.1.1 节末证过这个求积分的捷径.) 此外, 求第二个积分 (暂时忽略积分前的因子 3) 的最简单方法是要想到它等于半径为 1 个单位的半圆的面积 (平方单位), 即 $\pi/2$. 所以不用太多计算, 就可知整个答案为 $3\pi^2$, 因此半个面包圈的体积是 $3\pi^2$ 立方单位. 毫不意外, 我们刚才用的方法称为*壳法* (也称为*柱壳法*).

29.1.3 总结和变式

到目前为止, 我们已经知道了如何在半圆的例子中应用圆盘法和壳法. 相同的方法也可应用在由曲线、x 轴和两个垂线围成的区域, 如图 29-10 所示.

与半圆例子中的论证一样, 我们可以得到下面的原理.

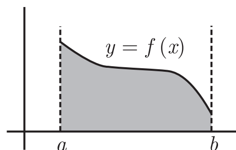


图 29-10

- 若将曲线 $y=f(x)$ 下方、$x=a$ 和 $x=b$ 之间围成的区域绕 x 轴旋转, 则可应用圆盘法, 其体积等于

$$\int_a^b \pi y^2\mathrm{d}x \text{ 立方单位.}$$

- 若将曲线 $y=f(x)$ 下方、$x=a$ 和 $x=b$ 之间围成的区域绕 y 轴旋转, 则可应用壳法, 其体积等于

$$\int_a^b 2\pi xy\mathrm{d}x \text{ 立方单位.}$$

能用心记住这些公式很多, 但若能了解如何求一般圆盘和壳体积并由此推导出这些公式就更好了. 当你遇到下面的变式之一时, 这点尤其有用:

(1) 要旋转的区域在曲线和 y 轴之间 (而不是 x 轴);

(2) 要旋转的区域在两曲线之间, 而不只是曲线下方到某个轴的区域;

(3) 旋转轴可能平行于 x 轴或 y 轴, 而不是轴本身.

对于这几种情况的任何组合, 你都可以通过选取小条并合理地旋转, 然后进行积分来求解. 在我们讨论如何实现之前, 首先重要的是如何确定应用圆盘法还是应用壳法. 注意, 应用圆盘法时, 小条绕平行于它们短边的轴旋转; 而应用壳法时, 小条绕垂直于短边的轴旋转. 也就是说, 在将区域切成小条之后,

- 若每个小条的短边**平行于**旋转轴, 运用**圆盘法**;
- 若每个小条的短边**垂直于**旋转轴时, 运用**壳法**.

有了这些知识, 我们就可以逐个来看下面的三个变式了.

29.1.4 变式 1：区域在曲线和 y 轴之间

如果区域在曲线和 y 轴之间, 你可能会取横着的小条, 短边在 y 轴上, 如图 29-11 所示.

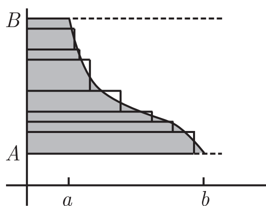


图 29-11

我们其实类似在求由曲线和 y 轴所围区域的面积, 见 16.4.3 节. 不管怎样, 若想求由该区域绕 y 轴旋转所得的立体的体积, 就应该用圆盘法, 因为小条的短边平行于 y 轴. 在 y 轴某处的一个小条, 其宽 $\mathrm{d}y$ 单位, 长 x 单位, 所得圆盘的体积为 $\pi x^2\mathrm{d}y$ 立方单位. 当对其求积分来求整个体积时, 要时刻注意积分的上下限对应 y 轴上的点, 而不是 x 轴上的点, 因为积分是关于 y 的 (因为 $\mathrm{d}y$). 具体说, 积分应从 A 到 B, 而不是 a 到 b, 所以我们要求的体积是 $\int_A^B \pi x^2\mathrm{d}y$.

还有另一种方法. 把头偏向右肩观察图 29-11, 此时 y 轴变成了水平的, 一切都倒置了, 试着想象一下, 如果纸是透明的, 且从反面看图 (头还偏着的) 会发生什么. 现在 y 轴和 x 轴交换了位置! 这就表明, 交换 x 和 y 变量是可行的, 假定你仍用 y 轴上的点做积分范围. 其实, 如果对 29.1.3 节的公式 $V=\int_a^b \pi y^2\mathrm{d}x$ 进行这个变换, 我们可以看到曲线到 y 轴的区域绕 y 轴旋转所得体积是 $\int_A^B \pi x^2\mathrm{d}y$, 这与我们之前所求一致.

如果上面的区域绕 x 轴而不是绕 y 轴旋转呢? 只需用 29.1.3 节的壳公式$V=\int_a^b 2\pi xy\mathrm{d}x$, 可知我们想要的体积是 $\int_A^B 2\pi yx\mathrm{d}y$. 这是有道理的, 因为将小条关于 x 轴旋转将得到一个厚 $\mathrm{d}y$、高 x, 半径为 y 个单位的壳. 你应该画一下, 看看这个小条展开成薄片时会是什么样. 这个薄片近似为一个立方体, 计算它的体积, 并确实得到 $2\pi yx\mathrm{d}y$. 综上, 所得的准则如下.

> 若区域在曲线和 y 轴之间, 交换 x 和 y.

如往常一样, 任意画一个小条, 旋转, 计算所得体积, 积分是最可信赖的方法. 上述准则只是个指导.

下面是变式 1 的例子. 令 R 为曲线 $y=\sqrt{x}$、$y=2$ 和 y 轴之间的区域, 如图 29-12 所示. 我们来计算 R 关于 y 轴和 x 轴的旋转体的体积. 在第一种情况中, 我们用圆盘法, 因为区域在曲线和 y 轴之间, 且关于相同的 y 轴旋转, 则体积为

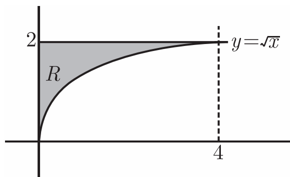


图 29-12

$$\int_0^2 \pi x^2 \mathrm{d}y.$$

因为 $y=\sqrt{x}$, 我们有 $x=y^2$, 所以 $x^2=y^4$. 因此, 体积是

$$\int_0^2 \pi x^2 \mathrm{d}y = \pi\int_0^2 y^4 \mathrm{d}y = \left.\frac{\pi y^5}{5}\right|_0^2 = \frac{32\pi}{5}$$

立方单位. 第二种情况, R 关于 x 轴的旋转体的体积用壳法求, 其体积为

$$\int_0^2 2\pi yx \mathrm{d}y = 2\pi\int_0^2 y^3 \mathrm{d}y.$$

因为 $yx=y\times y^2=y^3$, 可验证得出为 8π 立方单位. 要确定你能画出该例两种情况下的小条, 并验证上面公式的正确性. 同时也要注意积分必须从 0 到 2, 而不是从 0 到 4: 毕竟, 积分是关于 y(而非 x) 的, 而 y 的取值范围为 $[0,2]$, 如图所示.

29.1.5　变式 2: 两曲线间的区域

如果要旋转的区域介于两曲线之间, 那么我们面对的将与 16.4.2 节求两曲线间面积一样的情形. 一般方法是取顶部曲线下方到旋转轴的区域进行旋转, 得到一个较大的立体; 再取底部曲线下方到旋转轴的区域进行旋转, 得到较小的立体, 从较大立体中去掉较小立体, 得到的就是所需的立体. 考虑图 29-13 中的三个区域.

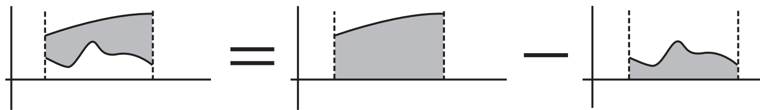


图　29-13

我们要旋转的区域见左图, 它是顶部曲线下方到 x 轴 (中图) 的区域和底部曲线下方到 x 轴区域 (右图) 之差. 不管是关于 x 轴还是关于 y 轴旋转, 我们所求区域的旋转体体积都等于较大区域旋转体体积与较小区域旋转体体积之差. 例如, 若该区域绕 x 轴旋转, 则得到一个类似截去头的圆锥, 且它的中间从左到右有个样子奇特的洞. 该立体是实体 (没有洞) 与洞的差, 如图 29-14 所示.

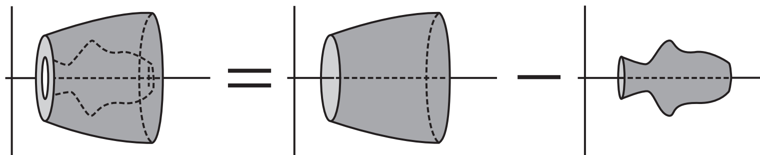


图　29-14

因此, 我们推出了如下结论.

若区域在两曲线之间, 求相应的旋转体体积之差.

来看一个具体的例子. 考虑如图 29-15 所示的两曲线 $y = 2x^3$ 和 $y = x^4$ 之间的有限区域. 该区域绕 x 轴旋转所得的立体的体积是多少?

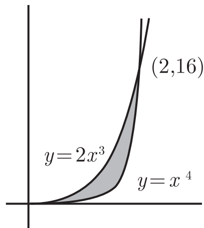


图 29-15

欲求交点, 需令 $2x^3 = x^4$. 可得 $x = 0$ 或 $x = 2$. 所以两条曲线的交点为原点和 $(2, 16)$, 如图所示. 因此, 我们要考虑的 x 区间是 $[0, 2]$. 现在, 对应 x 的区间, 曲线 $y = 2x^3$ 在曲线 $y = x^4$ 上方, 所以我们要求的体积是 $y_1 = 2x^3$ 的旋转体体积减去 $y_2 = x^4$ 的旋转体体积. 注意, 我们用 y_1 和 y_2 来取代 y 有利于区分二者. 现在对两个曲线分别用圆盘法, 则我们所求的体积是

$$\int_0^2 \pi y_1^2 \mathrm{d}x - \int_0^2 \pi y_2^2 \mathrm{d}x = \pi \int_0^2 (2x^3)^2 \mathrm{d}x - \pi \int_0^2 (x^4)^2 \mathrm{d}x.$$

你可以计算结果, 并验证答案为 $1024\pi/63$ 立方单位.

若是此区域关于 y 轴旋转呢? 我们在前面求得了两曲线间的面积, 但并没有特意倾向某个轴或其他轴, 所以应该能够用圆盘法或壳法来求解. 我们分别用两种方法来实现. 首先用圆盘法. 假设我们将该区域切割成短边平行于 y 轴的小条, 如图 29-16 所示. 所求的体积是 $y = x^4$ 和 $y = 2x^3$ 旋转体体积之差. 在这两个体积中, 第一个大于第二个, 因为 x^4 在 $2x^3$ 的右边, 所以我们令 $x_1 = y^{1/4}$, $x_2 = (y/2)^{1/3}$. 运用圆盘法, 将 x 和 y 互换 (如变式 1), 并在 $y = 0$ 和 $y = 16$ 间积分 (不是从 0 到 2), 可知所求体积为

$$\begin{aligned}\int_0^{16} \pi x_1^2 \mathrm{d}y - \int_0^{16} \pi x_2^2 \mathrm{d}y &= \pi \int_0^{16} (y^{1/4})^2 \mathrm{d}y - \pi \int_0^{16} ((y/2)^{1/3})^2 \mathrm{d}y \\ &= \pi \int_0^{16} y^{1/2} \mathrm{d}y - 2^{-2/3}\pi \int_0^{16} y^{2/3} \mathrm{d}y.\end{aligned}$$

稍作几步运算后, 可知结果为 $64\pi/15$ 立方单位. 你可以练习算一下.

下面用壳法来求相同的体积. 这次, 我们垂直切割该区域, 如图 29-17 所示. 由于 $y_1 = 2x^3$ 在 $y_2 = x^4$ 之上, 所以取两体积之差, 得到

$$\begin{aligned}&\int_0^2 2\pi x y_1 \mathrm{d}x - \int_0^2 2\pi x y_2 \mathrm{d}x \\ &= 2\pi \int_0^2 2x^4 \mathrm{d}x - 2\pi \int_0^2 x^5 \mathrm{d}x,\end{aligned}$$

结果为 $64\pi/15$ 立方单位, 这与用圆盘法所求结果一样. 这是当然! 注意, 我们用圆盘法时, 把所求立体看作是一个中间挖掉另一个碗的碗状物, 而用壳法时, 所求立体更像是一个中间去掉一个更小盆的盆状物. 你应该画图来看看具体情况.

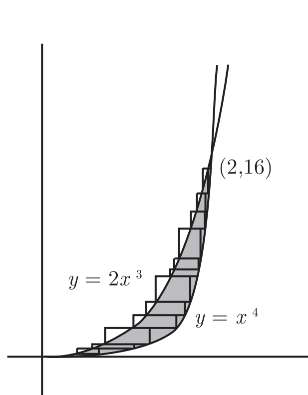


图　29-16

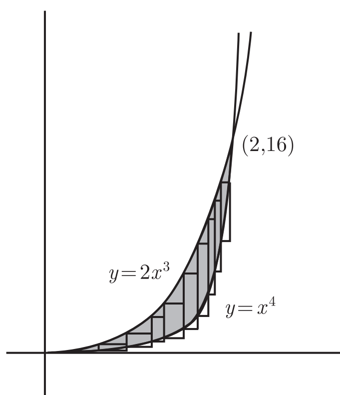


图　29-17

这个变式同样适用于没有延伸到坐标轴的区域. 例如, 假设我们要求曲线 $y = 1 + \sqrt{25 - x^2}$ 和直线 $y = 1$ 之间的区域绕 x 轴旋转所得的旋转体体积. 注意, 曲线是中心在 $(0, 1)$、半径为 5 个单位的圆 $x^2 + (y-1)^2 = 25$ 的上半部分, 其涉及区域如图 29-18 所示.

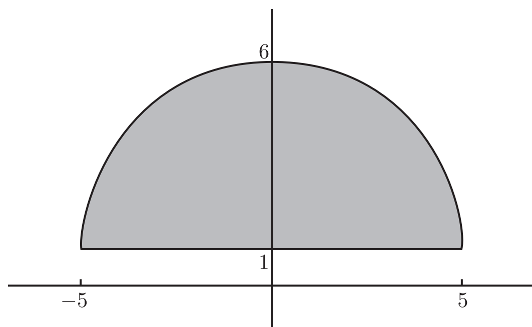


图　29-18

当我们将该区域绕 x 轴旋转时, 得到一个类似串珠的形状 —— 一个中心有洞的球状立体. 它的体积是多少呢? 你可以用关于变量 y 的壳法把它作为练习吧①. 另一个可行的方法是圆盘法. 我们应该将该区域看作曲线 $y_1 = 1 + \sqrt{25 - x^2}$ 和 $y_2 = 1$ 之间的区域, 所以体积为

$$\int_{-5}^{5} \pi(1 + \sqrt{25 - x^2})^2 \mathrm{d}x - \int_{-5}^{5} \pi(1)^2 \mathrm{d}x.$$

第二个积分是 10π, 恰巧是高 10 个单位、底面半径 1 个单位的圆柱体体积 —— 正是串珠中间空心部分. 第一个积分留给你计算, 记住 $\int_{-5}^{5} \sqrt{25 - x^2}\mathrm{d}x$ 比你想得要简

① 这个练习必须小心, 因为该区域并不是关于 y 轴的曲线下方部分. 最好是先算出半圆的右半边绕 x 轴旋转的体积, 然后令结果乘 2.

单 —— 不需计算, 因为它就是半径为 5 个单位的半圆面积. 不管怎样, 你应该验证答案为 $25\pi^2 + 500\pi/3$ 立方单位.

29.1.6 变式 3: 绕平行于坐标轴的轴旋转

最后, 我们来看一下如何处理轴为 $x = h$ 或 $y = h$ 的旋转体, 其中 h 不必一定等于 0. 我们从 $y = h$ 开始, 它平行于 x 轴, 高为 h. 假设我们令曲线 $y = f(x)$、直线 $y = h$、$x = a$ 和 $x = b$ 间的区域绕直线 $y = h$ 旋转, 如图 29-19 所示.

如图所示的小条宽为 $\mathrm{d}x$, 但高不是 y, 而是 $y - h$. 在图中, h 显示为正数, 所以显然 $y - h$ 小于 y—— 事实上也是如此. 若碰巧 h 为负, 则小条的高大于 y —— 显然这时 $y - h$ 大于 y, 因为 h 为负! 不考虑 h 的符号, 我们看到小条的高为 $y - h$, 所以相应的圆盘体积是 $\pi(y - h)^2\mathrm{d}x$, 整个旋转体的体积是 $\int_a^b \pi(y - h)^2\mathrm{d}x$.

事实上, 这个公式与正规圆盘法的唯一区别是: y 被 $(y - h)$ 代换了. 如 1.3 节所述, 这个变化转换了标准图像, 使得区域位在 x 轴上方 h 单位的地方 (若 h 为负则在 x 轴下方). 该变化的唯一问题是, 直线 $y = h$ 可能会在曲线上方, 图 29-20 所示. 在这种情况下, 小条的高度是 $h - y$, 而不是 $y - h$. 这对圆盘法没有实质影响, 因为是取高的平方, 但对此加以小心总是好的. 壳法则另当别论了.

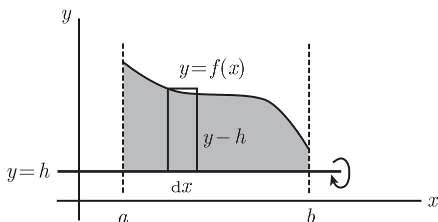


图 29-19

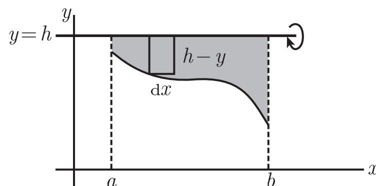


图 29-20

假设我们欲求由如图 29-21 所示区域绕轴 $x = h$ 旋转得到的立体体积. 这里我们要用壳法, 因为小条的短边垂直于旋转轴. 一个壳高 y、厚 $\mathrm{d}x$ 单位, 不过现在的半径是 $x - h$ 而非 x 单位. 你可以验证壳的体积为 $2\pi(x - h)y\mathrm{d}x$, 所以总体积是 $\int_a^b 2\pi(x - h)y\mathrm{d}x$ 立方单位. 同样, 注意这个结论来源于 29.1.3 节的壳法标准公式, 只是其中的 x 代换为 $(x - h)$. 这个变换将标准图像向右平移了 h 单位 (包括旋转轴) —— 我们做的就是平移图像.

如果旋转轴在该区域右边呢? 考虑图 29-22. 现在壳的半径是 $h - x$ 单位, 而不是 $x - h$ 单位, 因为 h 大于积分区间 $[a, b]$ 内的所有 x. 所以这次旋转体的体积是 $\int_a^b 2\pi(h - x)y\mathrm{d}x$ 立方单位 (具体自行验证).

所以, 变式 3 的一般思想是:

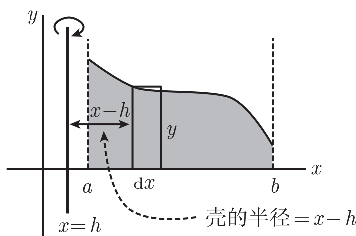


图　29-21

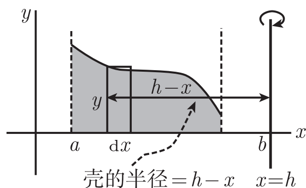


图　29-22

> 若旋转轴是 $x = h$, 用 $(x - h)$(若 $x < h$ 则用 $(h - x)$) 代换 x;
> 若旋转轴是 $y = h$, 用 $(y - h)$(若 $y < h$ 则用 $(h - y)$) 代换 y.

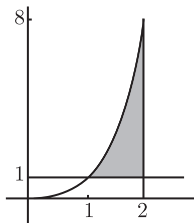


图　29-23

我们来看一些变式 3 的例子. 在这些例子中, 我们将讨论在曲线 $y = x^3$、直线 $x = 2$ 和 $y = 1$ 之间的区域, 如图 29-23 所示. (注意图中的 x 轴和 y 轴尺度不同, 因此这只是个粗略图.) 我们先来求该区域绕直线 $y = 1$ 旋转所得立体的体积. 为求该体积, 就要将 y 替换为 $y - 1$, 该图向下移 1 个单位. 因此体积是

$$\int_1^2 \pi(y-1)^2 \mathrm{d}x = \pi\int_1^2 (x^3-1)^2 \mathrm{d}x,$$

很容易算出它为 $163\pi/14$ 立方单位. 想一想能否通过求圆盘的体积来验证这个答案 (小条是垂直的).

若此区域绕直线 $x = 2$ 旋转呢? 这其实是变式 1 和变式 3 的组合, 由于旋转轴平行于 y 轴, 所以我们将交换 x 和 y, 并用 $(2 - x)$ 代换 x 来处理这个平移. 注意这里是 $(2 - x)$ 而不是 $(x - 2)$, 因为区域在直线 $x = 2$ 的左边. 同样, 积分应该从 1 到 8, 因为积分是关于 y 而不是关于 x 的. 因此体积为

$$\int_1^8 \pi(2-x)^2 \mathrm{d}y = \pi\int_1^8 (2-y^{1/3})^2 \mathrm{d}y,$$

化简后为 $8\pi/5$ 立方单位. 最好验证一下通过求圆盘体积也可求出该体积, 不过注意, 这次我们将区域切成了水平小条, 就像变式 1 一样.

若我们让此区域绕 $x = -3$ 旋转呢? 开始有点乱了. 若我们用垂直小条, 则需用壳法, 因为每个小条的短边垂直于旋转轴. 我们将组合使用变式 2 和变式 3. 垂直来看, 区域在两个曲线 $y_1 = x^3$(在顶部) 和 $y_2 = 1$(在底部) 之间. 同样, 壳法标准公式中的 x 要替换为 $(x + 3)$. 这意味着体积由

$$\int_1^2 2\pi(x+3)y_1 \mathrm{d}x - \int_1^2 2\pi(x+3)y_2 \mathrm{d}x = 2\pi\int_1^2 (x+3)x^3 \mathrm{d}x - 2\pi\int_1^2 (x+3)\mathrm{d}x$$

给出, 计算可得 $259\pi/10$ 立方单位.



我们来重复这个例子, 这次取水平小条. 现在我们要用圆盘法, 因为每个小条的短边平行于旋转轴. 我们需交换 x 和 y, 因为旋转轴是垂直的 (变式 1); 同样, 我们要把该区域看作平放在右曲线 $x_1 = 2$ 和左曲线 $x_2 = y^{1/3}$ 之间; 最后, 我们需将 x 代换为 $x+3$(变式 3), 意思是将 x_1 代换为 x_1+3 且将 x_2 代换为 x_2+3. 所以这个例子运用了以上三个变式! 标准圆盘体积是 $\pi y^2\mathrm{d}x$; 交换 x 和 y 可得 $\pi x^2\mathrm{d}y$; 用 $x+3$ 代换 x 得 $\pi(x+3)^2\mathrm{d}y$; 对该形式分别关于 x_1 和 x_2 从 1 到 8 积分, 取差. 可得体积为

$$\int_1^8 \pi(x_1+3)^2\mathrm{d}y - \int_1^8 \pi(x_2+3)^2\mathrm{d}y = \pi\int_1^8 (2+3)^2\mathrm{d}y - \pi\int_1^8 (y^{1/3}+3)^2\mathrm{d}y,$$

算出来仍为 259π / 10 立方单位. 至少我们得到了一样的答案! 同样, 最好你可以自己求圆盘体积.

至此, 我们已经有足够多关于旋转体体积的理论了, 要想掌握所有的变式就必须多做练习. 现在是时候讨论求更一般立体体积了.

29.2 一般立体体积

大多数立体不能通过平面区域绕平面内某轴旋转而形成. 例如, 一个棱锥没有曲面, 所以无论你怎么看, 它都不是旋转体. 求类似立体体积的一个方法是**切片法**, 这是推广了 29.1.1 节的圆盘法.

把立体想象成一种蔬菜, 比如黄瓜或南瓜. 将它放在案板上切成薄的平行的切片. 这些切片的大小不会全部相同, 甚至一个切片的两面也不一样. 例如黄瓜, 靠近端部的切片会有点斜. 另一方面, 若切片很薄, 则它的两面会很接近. 所以我们将取其中一面的面积乘上切片的厚度来近似切片的体积 —— 取哪面都没关系. 然后我们将把所有切片的体积加起来, 求切片厚度趋于 0 的极限.

在实践中, 这个过程有些复杂. 事实上, 有很多方法来切割立体. 例如, 若切平放的黄瓜, 则得到盘状的薄切片; 若切竖放的黄瓜, 虽然较难, 不过还是可行的, 你会得到大小不同的椭圆形切片. 或者, 将黄瓜倾斜一个角度, 切得更小的椭圆形.

基本上, 你的选择是：选择一个轴, 它不必穿过立体. 所有的切片将垂直于这个轴. 一旦选定了轴, 后续的思路就清晰了：求得每个垂直于该轴的切片的横截面面积. 不同的切片有不同的面积. 所以, 要在轴上选择一个原点和正方向, 然后算出穿过 x 的切片的横截面面积, 其中 x 是轴上的任意一点. 最后一步是用面积乘厚度 $\mathrm{d}x$ 来近似切片的体积, 然后积分. 这步相当于把所有切片的体积加起来, 同时取切片最大厚度趋于 0 的极限. 综上, 解题思路是：

(1) 选定一个轴;

(2) 求轴上点 x 处的切片横截面面积, 称该面积为 $A(x)$ 平方单位;

(3) 若 V 为立体的体积 (立方单位), 我们有

$$V = \int_a^b A(x)\mathrm{d}x,$$

其中 $[a, b]$ 是完全覆盖立体的 x 的取值范围.

相信我, 你一定要选一个使横截面越简单越好的轴. 最好能确保横截面都很相似, 也就是它们互为不同大小的副本. 不过, 这个可能也不是总有.

让我们用上述方法求一个 "广义" 锥体的体积. 这里的意思是, 在平面上有面积为 A 平方单位的某个形状, 平面上方一定距离处有一顶点 P, 如图 29-24 所示. 现在, 我们做从平面形状边上的每个点到 P 的线段, 这就得到一个底为起始形状的曲面. 我们要讨论的立体就是被填充了的曲面, 或者说是曲面的内部. 图 29-25 为大概的曲面框架图.

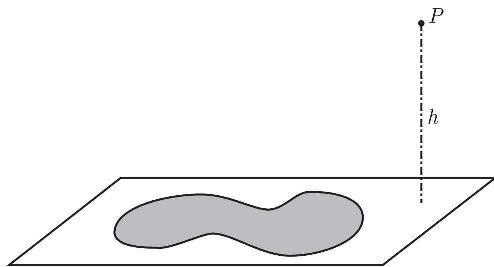


图 29-24

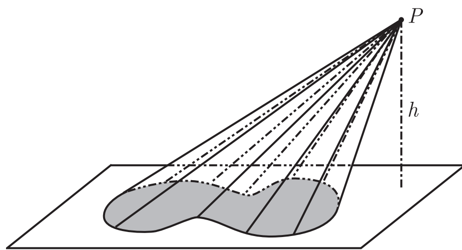


图 29-25

例如, 若底为圆且点 P 在圆心的正上方, 则我们得到一个通常的圆锥. 若底为正方形且点 P 在正方形中心 (即正方形对角线的交点) 的正上方, 则我们得到一个四棱锥. 你可以想象一下, 什么样的底和顶点 P 可以得到规则的锥体或斜锥体 (就像一顶奇怪的帽子, 类似巫师帽, 但不是直的). 结果表明, 与求立体体积相关的仅有的量是底面积 —— A 平方单位, 以及点 P 到平面的垂直距离 —— h 单位 (图 29-25 已标出).

那么, 如何求体积呢? 首先要选一个轴. P 似乎是一个特殊的点, 所以我们选择的直线或许应该穿过 P. 那其他点呢? 你可以进行各种尝试, 但唯一有用的是让直线垂直于底所在的平面. 我们也把轴的原点设在 P, 正方向向下, 这会使计算更容易. (看起来有点怪, 但谁说正方向不可以向下呢. 毕竟, 广义锥体也可能顶点在下, 此时向上是正方向.) 我们来看若选择轴上的点 x 并取穿过 x 的垂直切片会怎样 (参见图 29-26).

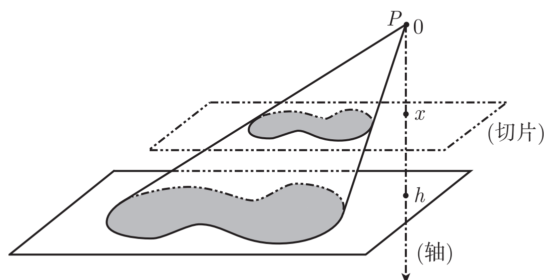


图 29-26

横截面是原底的一个较小副本. 用数学语言讲就是横截面与底相似. 现在我们要求横截面的面积. 为此, 我们选取底的边上任意一点并连接到 P. 这条线在广义锥体的边上, 且穿过较小截面上的相应点. 我们选的点最好能使得直线位于图像的右边缘, 当然我们也可以选底的边上任意点. 我们还要画一些垂线段, 如图 29-27 所示.

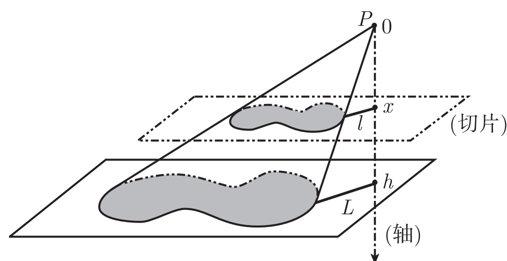


图 29-27

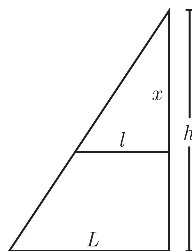


图 29-28

我在上图标出了垂线的长度. 垂线形成的三角形如图 29-28 所示.

运用相似三角形, 我们可知

$$\frac{x}{l} = \frac{h}{L},$$

这意味着 $l = xL/h$. 我们验证一下这个方程. 若 $x = 0$, 则切片过锥体的顶点 (P) 且 l 应该为 0, 而它就是 0. 另一方面, 若 $x = h$, 则切片是底, 且横截面不是底的较小副本 —— 它就是底. 所以, 这时 l 理所当然应该等于 L. 事实的确如此.

现在我们来看底和横截面, 其中画出了长度为 L 和 l 的线段, 如图 29-29 所示.

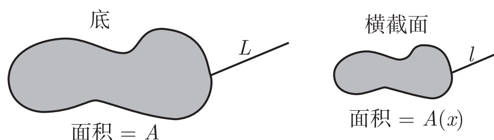


图 29-29

在这两幅图中, 包括线段在内都是相似的 —— 一个是另一个的精确放大. 这里有一个相似性的重要原理. 假设我们有两个相似图形, 且已知两个图形中对应线段的长度. 当我们将一个图形放大到与另一个图形一样大小时, 两条线段应该严格匹配. 那么, 两个图形的面积之比就是对应线段长度之比的平方. 例如我们取两个正方形的瓷砖, 其中一个边长是另一个边长的 3 倍, 则大瓷砖的面积是小瓷砖面积的 9 倍. 回到上图, 底的面积是 A 平方单位, 横截面的面积是 $A(x)$ 平方单位. 因此, 面积之比是对应线段长度之比的平方, 在本例中长度是 L 和 l, 则

$$\frac{A}{A(x)} = (\frac{L}{l})^2.$$

化简并运用前面 l 的表达式, 可得

$$A(x) = \frac{Al^2}{L^2} = \frac{A}{L^2} \cdot \left(\frac{xL}{h}\right)^2 = \frac{Ax^2}{h^2}$$

同样来验证一下：若 $x = 0$, 横截面为点 P, 此时横截面没有面积. 得到验证, 因为 $A(0) = A \times 0^2/h^2 = 0$. 那 $x = h$ 时呢? 这时我们讨论的是底, 所以横截面面积应该为 A 平方单位. 没问题：$A(h) = A \times h^2/h^2 = A$.

最后, 我们可以做积分了! 唯一的问题是 x 的范围是多少. 我们可知, $x = 0$ 是顶, $x = h$ 是底, 这就是 x 的正确取值范围. 所以

$$V = \int_0^h A(x)\mathrm{d}x = \int_0^h \frac{Ax^2}{h^2}\mathrm{d}x = \frac{A}{h^2}\int_0^h x^2\mathrm{d}x = \frac{A}{h^2} \cdot \frac{h^3}{3} = \frac{1}{3}Ah$$

立方单位.

好了, 我们已求得任意棱锥或类圆锥体的体积公式. 例如讨论过的正圆锥, 体积是 $\frac{1}{3}\pi r^2 h$ 立方单位, 正是根据上面公式由 $A = \pi r^2$ 求得的结果. 对正四棱锥也一样有效, 其体积为 $\frac{1}{3}l^2 h$ 立方单位(其中底边长 l 单位), 因为此时底面积是 $A = l^2$.

我们再来看一个例子. 取在 $x = 0$ 和 $x = \frac{1}{2}$ 之间的曲线 $y = \mathrm{e}^x$, 并考虑曲线和 x 轴之间的区域. 如图 29-30 所示. 假设有一个形状怪异的立体位于上述平面的上方, 并延伸出纸面, 它的底就是上图的阴影区域. 该立体的形状是：若沿平行于 y 轴的任何直线竖直向下切, 则其横截面是一个矩形, 长边位于上图的底上, 短边为长边一半. 将图稍微倾斜一下来看透视图, 这些横截面的样子如图 29-31 所示.

该立体的体积是什么? 我们先来选轴. x 轴怎样? 似乎有道理, 因为我们知道垂直于该轴的横截面是什么样的. 我们已经有了原点和正方向, 那就以它们为准吧. 在轴上的 x 点, 垂线段长度为 e^x 单位. 这是矩形长边, 所以短边长度为 $\frac{1}{2}\mathrm{e}^x$ 单位(要知道, 短边是长边的一半). 因此矩形的面积为

$$A(x) = \mathrm{e}^x \times \frac{1}{2}\mathrm{e}^x = \frac{1}{2}\mathrm{e}^{2x}$$

平方单位. 故体积是

$$V=\int_0^{1/2}A(x)\mathrm{d}x=\frac{1}{2}\int_0^{1/2}\mathrm{e}^{2x}\mathrm{d}x=\frac{1}{2}\frac{\mathrm{e}^{2x}}{2}\bigg|_0^{1/2}=\frac{1}{4}(\mathrm{e}-1)\text{ 立方单位.}$$

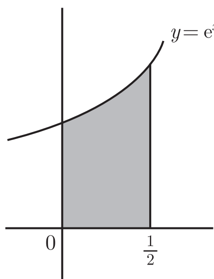


图　29-30

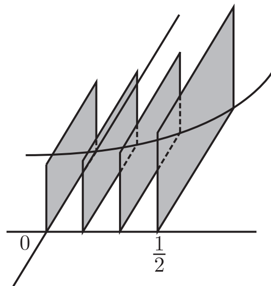


图　29-31

29.3　弧　　长

对于某函数 f, 我们有 $y=f(x)$ 的图像, 其中 x 的取值范围为 a 到 b. 取一段细绳, 顺着曲线摆放, 标记两端, 然后细绳离开纸面, 拉直并测量两个标记点之间的长度. 你怎么计算曲线长度呢? 这个长度称为曲线弧长, 我们希望找到一个求弧长的公式. 其策略是先得到某个基本表达式, 然后加以修改, 得到一些有用的公式.

我们来看介于 x 和 $x+\mathrm{d}x$ 之间的一小段曲线, 如图 29-32 所示.

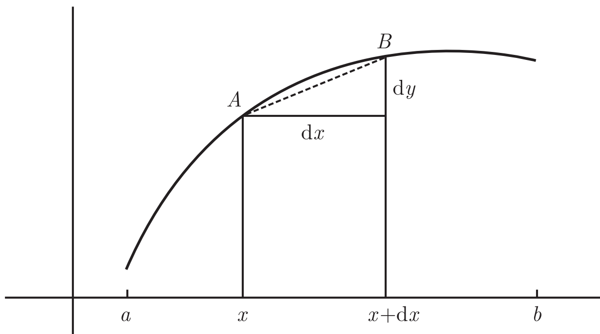


图　29-32

我们用虚线段 AB 的长度来近似 A 和 B 之间的曲线长度. A 与 B 越接近, 近似程度就越好. 根据勾股定理, AB 的长度是 $\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}$ 单位. 现在, 我们只需对很多小线段重复该过程, 就形成了对曲线的近似. 像往常一样, 积分关注连加和极限部分, 要小心. 若只在小段长度 $\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}$ 前加一个积分号, 会得到

$$\text{弧长}=\int_?^?\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}.$$

问题是, 这个积分没有任何意义! 我们需要关于变量的积分. 幸运的是, 我们可以针对各种情形来调整上面的公式, 从而得到有意义的结果. 例如, 你可以将因子 $(\mathrm{d}x)^2$ 移到根号外面, 将小段长度表示为 $\sqrt{1+(\mathrm{d}y/\mathrm{d}x)^2}\mathrm{d}x$ 单位. 这看起来更可行. (这个变动其实需要证明, 不过其细节超出了本书的范围.) 不管怎样, 在下面的每个例子中, 我们将讨论如何调整上面的基本公式来得到合乎情理的弧长公式.

(1) 若 $y=f(x)$, 且 x 在 a 到 b 间取值, 则在上述被积函数中取因子 $(\mathrm{d}x)^2$ (如前所述) 并将其移到根号外面得

$$\boxed{\text{弧长} = \int_a^b \sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x} \text{ (标准形式)}.$$

将其写成关于 f 的形式为

$$\text{弧长} = \int_a^b \sqrt{1+(f'(x))^2}\mathrm{d}x.$$

(2) 假设给定关于 y 的 x. 若 $x=g(y)$ 且 y 在 A 到 B 间取值, 则取因子 $(\mathrm{d}y)^2$(或者, 交换上面方框中公式的 x 和 y) 得

$$\boxed{\text{弧长} = \int_A^B \sqrt{1+\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2}\mathrm{d}y} \text{ (关于 } y\text{)},$$

也可写为

$$\text{弧长} = \int_A^B \sqrt{1+(g'(y))^2}\mathrm{d}y.$$

(3) 参数形式呢? 这意味着 x 和 y 是关于参数 t 的函数, t 在 t_0 到 t_1 间取值. (参数方程参见 27.1 节.) 我们将量 $(\mathrm{d}x)^2$ 看作 $(\mathrm{d}x/\mathrm{d}t)^2(\mathrm{d}t)^2$; y 同理. 然后可将 $(\mathrm{d}t)^2$ 移到根号外面, 得到一个有用的公式:

$$\boxed{\text{弧长} = \int_{t_0}^{t_1} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\mathrm{d}t} \text{ (参数版)}.$$

(4) 最后, 这个公式的特殊情况发生在极坐标情形. 特别地, 在 27.2.4 节, 我们讨论了如何求曲线 $r=f(\theta)$ 内部的面积, 其中 θ 的取值范围为 θ_0 到 θ_1, 现在我们来求相同曲线的弧长. 我们知道 $x=r\cos(\theta)$, $y=r\sin(\theta)$, 所以用 $f(\theta)$ 代换 r, 可得 $x=f(\theta)\cos(\theta)$, $y=f(\theta)\sin(\theta)$. 这里的 θ 即为参数, 所以我们可以使用参数版的弧长公式 (t 代换为 θ). 我们需知道 $\mathrm{d}x/\mathrm{d}\theta$ 和 $\mathrm{d}y/\mathrm{d}\theta$ 是什么. 根据乘积法则,

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = f'(\theta)\cos(\theta) - f(\theta)\sin(\theta)$$

和

$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = f'(\theta)\sin(\theta) + f(\theta)\cos(\theta).$$

现在需对这两个式子取平方并相加. 试一下吧! 你会发现有些项消掉了; 另外, $\sin^2(\theta)+\cos^2(\theta)$ 项可用 1 代换. 综上, 可得到公式



$$\boxed{\text{弧长}=\int_{\theta_0}^{\theta_1}\sqrt{(f(\theta))^2+(f'(\theta))^2}\mathrm{d}\theta}\quad (\text{极坐标函数}, r=f(\theta)).$$

顺便说一下, 你应该在表示弧长时带上单位.



我们来看一些例子. 假设要求曲线 $y=\ln(x)$ 的弧长, 其中 x 的取值范围为 $\sqrt{3}$ 到 $\sqrt{15}$. 我们用前面的第一个公式可得

$$\begin{aligned}\text{弧长}&=\int_{\sqrt{3}}^{\sqrt{15}}\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x=\int_{\sqrt{3}}^{\sqrt{15}}\sqrt{1+\left(\frac{1}{x}\right)^2}\mathrm{d}x\\&=\int_{\sqrt{3}}^{\sqrt{15}}\frac{\sqrt{x^2+1}}{x}\mathrm{d}x.\end{aligned}$$



这其实是个非常难的积分. 你一定要练习一下. 如果卡住了, 对策是：从一个合适的三角换元开始. 若做对了, 积分对应的不定积分为 $\int\sec^3(\theta)/\tan(\theta)\mathrm{d}\theta$. 要求它, 可将分子表示为 $\sec(\theta)(1+\tan^2(\theta))$, 将原积分分成两个积分, 再用第 19 章的方法求解. 可验证所得弧长为 $2+\ln(3)-\dfrac{1}{2}\ln(5)$ 单位.



若弧长是由参数 $x=3t^2-12t+4$ 和 $y=8\sqrt{2}t^{3/2}$ 表述的, 其中 t 在 3 到 5 间取值, 该怎么求呢？我们需用参数版的公式. 事实上, $\mathrm{d}x/\mathrm{d}t=6t-12$, $\mathrm{d}y/\mathrm{d}t=12\sqrt{2}t^{1/2}$, 故

$$\text{弧长}=\int_3^5\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\mathrm{d}t=\int_3^5\sqrt{(6t-12)^2+(12\sqrt{2}t^{1/2})^2}\mathrm{d}t.$$

现在, 我们来看被积函数的最里面部分, 其中有一个因子 6^2 可被提出来, 得

$$\begin{aligned}(6t-12)^2+(12\sqrt{2}t^{1/2})^2&=6^2((t-2)^2+(2\sqrt{2}t^{1/2})^2)\\&=36(t^2-4t+4+8t)=36(t+2)^2.\end{aligned}$$



现在将这个结果带入被积函数并作积分, 可得弧长为 72 单位. 这就是一件简单的事了, 细节留给你完成!

参数化和速率

在讨论求表面积之前, 我还想谈谈关于参数坐标系下弧长公式的一点事. 假设一只蚂蚁 (这次不是蜗牛!) 绕一个平地爬行, 我定义在时间 t 秒处的蚂蚁位置是 $(x(t),y(t))$. 那么, 蚂蚁在时间 t 的速率是多少？我们知道速度是位移关于时间的导数. 因此蚂蚁在 x 方向的速度是 $\mathrm{d}x/\mathrm{d}t$, 在 y 方向的速度是 $\mathrm{d}y/\mathrm{d}t$. 它的实际速率

需涉及这两个速度. 其实, 根据毕达哥拉斯定理, 我们应该有[①]：

$$\boxed{速率 = \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}.}$$

嘿, 这是在参数情况下求弧长时所积分的量啊! 确实, 要求蚂蚁爬过的总距离, 需要对它的速率求积分. 因此, 现在弧长公式中的被积函数有了意义, 至少在参数情形下是有意义的：它是质点在曲线上移动的瞬时速率, 就像参数所描述的一样.

考虑前一节末的例子, 其中 $x = 3t^2 - 12t + 4$, $y = 8\sqrt{2}t^{3/2}$. 根据前面的探讨,

$$速率 = \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} = \sqrt{36(t+2)^2} = 6(t+2).$$

其中答案以单位每秒表示 (假设 t 的单位为秒). 这意味着在时间 $t = 3$ 处, 质点 (此时位于 $(x(t), y(t))$ 的速率是 $6(3+2) = 30$ 单位每秒; 而在时间 $t = 5$ 处 (速率稍快一点), 为 $6(5+2) = 42$ 单位每秒.

在 27.1 节, 我们探讨了参数方程 $x = 3\cos(t)$ 和 $y = 3\sin(t)$ $(0 \leqslant t < 2\pi)$ 所描述的中心在原点、半径为 3 的圆. 由这些方程所描述的运动的质点速率为

$$\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} = \sqrt{(-3\sin(t))^2 + (3\cos(t))^2} = \sqrt{9} = 3,$$

因为 $\sin^2(t) + \cos^2(t) = 1$. 这意味着质点以恒定的速率 3 单位每秒绕圆运动 (当然是逆时针方向). 另一方面, 我们也探讨了 $x = 3\cos(2t)$ 和 $y = 3\sin(2t)$ $(0 \leqslant t < \pi)$ 所描述的相同的圆, 这时的速率是

$$\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} = \sqrt{(-6\sin(2t))^2 + (6\cos(2t))^2} = \sqrt{36} = 6,$$

所以这个新的参数方程的质点确实以两倍于原质点的速率绕相同的圆运动.

29.4　旋转体的表面积

本章最后要讨论的问题, 是如何求由曲线绕某轴旋转所得表面的表面积. 我们采用的方法结合了求弧长和体积的方法. 我们从将曲线切割成小段弧开始, 然后关注绕轴旋转其中一段弧时的情况. 假设绕 x 轴旋转. 当旋转这一小段弧时会发生什么呢？我们得到了一个环, 但它的边是弯的. 若环的宽足够小, 我们应该能用直边环来近似它. 我们从割线段近似弧开始, 如 29.3 节的做法. 可见, 割线的长为 $\sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2}$ 单位. 当我们用该割线代替弧段旋转时, 得到了一个直边环, 如图 29-33 所示.

① 这里进入了向量范畴, 它属于关于多变量微积分的书所涉及的内容.

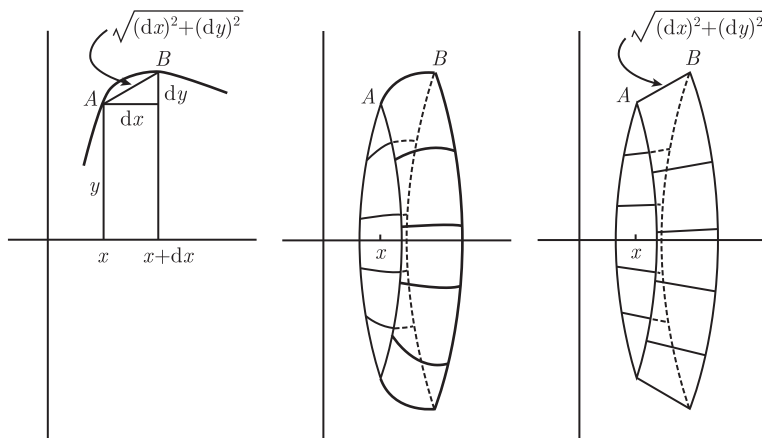


图 29-33

左边的图给出了一段曲线和近似割线; 中间的图给出了我们所求表面积的真实的曲边环; 右边的图给出了用于替换的近似环. 事实上, 我们可以更懒：环的边并不平行于 x 轴, 所以我们的环实际上是圆锥表面的一部分. 这类物体的表面积是可以计算的, 但比较麻烦. 因而我们进一步近似, 假想要讨论的是一个边长都相等的环, 不过这个环是一个圆柱形的, 如图 29-34 所示.

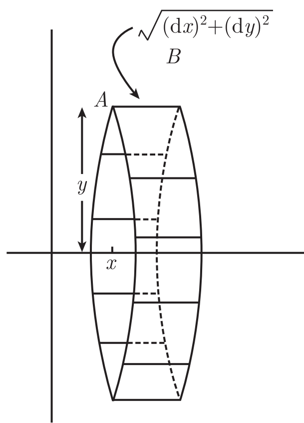


图 29-34

最终的结果是, 我们得到了一个半径为 y 单位、宽度为 $\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}$ 单位的圆柱形环, 因此它有表面积 $2\pi y\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}$ 平方单位(周长 $2\pi y$ 单位乘以宽度). 结果表明[①], 这个近似在极限中是可行的, 即将这些环的表面积加起来, 并令环的宽度趋于 0 的极限. 所以, 我们由此得到关于 x 轴旋转的原型公式：

$$\text{表面积}=\int_?^? 2\pi y\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}\ (\text{关于 } x \text{ 轴旋转}).$$

若旋转是关于 y 轴的, 则我们采用的环宽度不变, 但现在的半径是 x 而不是 y 单位, 所以关于 y 轴旋转的原型公式是

$$\text{表面积}=\int_?^? 2\pi x\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}\ (\text{关于 } y \text{ 轴旋转}).$$

你也可以参照体积的变式 1 (见 29.1.4 节), 将第一个原型公式中的 x 和 y 对换来得到该公式.

① 所牵涉的计算有些令人生厌. 如果你想算的话, 可运用半径为 r 和 R, 高为 h 单位的圆台的侧面积公式 $\pi(R+r)\sqrt{(R-r)^2+h^2}$ 平方单位.

不管怎样, 与弧长一样, 这些原型公式不能用于求任何表面积! 我们来看一下如何修改才能使用它们.

(1) 假设我们关于 x 轴旋转曲线 $y = f(x)$, 其中 x 取值范围为 a 到 b. 我们从第一个原型公式中的被积函数中提出因子 $(\mathrm{d}x)^2$ 并将其提到根号之外, 就像在讨论弧长时所做的一样, 得

$$\boxed{\text{表面积} = \int_a^b 2\pi y\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x} \text{ (关于 } x \text{ 轴).}$$

写成关于 f 的形式, 即为

$$\text{表面积} = \int_a^b 2\pi f(x)\sqrt{1+(f'(x))^2}\mathrm{d}x.$$

(2) 若我们关于 y 轴旋转相同的曲线, 可对另一个原型公式采用相同的处理, 得出

$$\boxed{\text{表面积} = \int_a^b 2\pi x\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x} \text{ (关于 } y \text{ 轴),}$$

或者写成关于 f 的形式为

$$\text{表面积} = \int_a^b 2\pi x\sqrt{1+(f'(x))^2}\mathrm{d}x.$$

(3) 当然也有参数形式. 若 x 和 y 是参数 t 的函数, 其中 t 的取值范围为 t_0 到 t_1, 则可推出下面的公式：

$$\boxed{\text{表面积} = \int_{t_0}^{t_1} 2\pi y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\mathrm{d}t} \text{ (参数版, 关于 } x \text{ 轴)}$$

和

$$\boxed{\text{表面积} = \int_{t_0}^{t_1} 2\pi x\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\mathrm{d}t} \text{ (参数版, 关于 } y \text{ 轴)}$$

所有这些表面积的单位都是平方单位.

来看一个例子. 若从 $x = 0$ 到 $x = \pi/2$ 的曲线 $y = \cos(x)$ 关于 x 轴旋转, 我们需应用第 1 种情形中的公式, 可知表面积为

$$\int_0^{\pi/2} 2\pi y\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x = 2\pi\int_0^{\pi/2}\cos(x)\sqrt{1+\sin^2(x)}\mathrm{d}x.$$

要求该积分, 首先令 $t = \sin(x)$, 然后用三角换元来求解新积分. 试一下, 算出来的表面积应为 $\pi(\sqrt{2}+\ln(1+\sqrt{2}))$ 平方单位.

另一方面, 在 $x = 0$ 和 $x = 2\sqrt{2}$ 间的抛物线 $y = x^2/2$ 绕 y 轴 (非 x 轴) 旋转

所得的表面积可用第 2 种情形中的公式求得. 由于 $\mathrm{d}y/\mathrm{d}x = x$, 表面积由

$$\int_0^{2\sqrt{2}} 2\pi x\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x = 2\pi\int_0^{2\sqrt{2}} x\sqrt{1+x^2}\mathrm{d}x$$

给出, 做换元 $t = 1 + x^2$ 后可算得 $52\pi/3$.



现在考虑以原点为中心、半径为 r 单位的上半圆. 参数形式为 $x = r\cos(\theta)$ 和 $y = r\sin(\theta)$, 其中 θ 的取值范围为 0 到 π (只取到 π 是为了只取上半圆). 我们关于 x 轴旋转该半圆, 将得到一个球, 它的表面积由情形 3 给出 (t 代换为 θ):

$$\int_0^{\pi} 2\pi y\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2}\mathrm{d}\theta = 2\pi\int_0^{\pi} r\sin(\theta)\sqrt{(-r\sin(\theta))^2+(r\cos(\theta))^2}\mathrm{d}\theta.$$

现在可利用 $\sin^2(\theta) + \cos^2(\theta) = 1$ 计算, 可得表面积为 $4\pi r^2$ 平方单位, 这就验证了传统公式.

最后, 我们来考虑类似于旋转体体积变式 3 的表面积 (见 29.1.6 节). 若旋转轴不是 x 轴, 而是直线 $y = h$(平行于 x 轴), 则圆柱形环的半径是 $y - h$ 单位 (而非 y 单位), 所以情形 1 中的公式需要做适当修改:

$$\text{表面积} = \int_a^b 2\pi(y-h)\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x \quad (\text{关于}y = h).$$

(其实, 若曲线在直线 $y = h$ 下方, 最好用 $h - y$ 代替 $y - h$, 否则将得到表面积为负的答案!) 同样, 你不能单纯地学习上面的公式, 而应理解如何由已知推出该公式. 事实上, 你现在应该能够对前面所有的公式进行适当改动, 正确处理关于 $y = h$ 或 $x = h$ 的旋转.

第 30 章 微分方程

微分方程就是包含导数的方程, 它们对于描述现实世界中量的变化非常有用. 例如, 若想了解种群增长快慢, 或是还清学生贷款的快慢, 都可以使用微分方程模拟对应情形从而得出令人满意的答案. 在本书的这最后一章, 我们将讨论如何求解特定类型的微分方程. 下面是我们将讨论的内容:

- 微分方程导论;
- 可分离变量的一阶微分方程;
- 一阶线性微分方程;
- 一阶和二阶常系数微分方程;
- 微分方程建模.

30.1 微分方程导论

早在 9.6 节讨论指数增长和衰退时, 我们就已经见过微分方程的例子. 当时的方程是

$$\frac{\mathrm{d}y}{\mathrm{d}x} = ky,$$

其中 k 是确定的常数, 并断言唯一解的形式为 $y = A\mathrm{e}^{kx}$, A 为常数. 我们将在后面的 30.2 节证明这个断言. 我们不应该对这个突然出现的形如 A 的常数感到意外. 毕竟, 原方程包含一个导数, 消去导数的唯一方法就是对其积分, 而积分会引入一个未知常数 (回想 $+C$).

方程 $\mathrm{d}y/\mathrm{d}x = ky$ 是一个一阶微分方程的例子, 因为方程中只有一个一阶导. 一般地, 一个微分方程的阶是其所包含的最高阶导数的阶. 例如, 这个复杂的方程

$$x^2\frac{\mathrm{d}^4y}{\mathrm{d}x^4} + \sin(x)\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^7 + \mathrm{e}^x y = \tan(x)$$

是一个四阶微分方程, 因为它包含一个四阶导数, 没有五阶或更高阶导数.



现在考虑本节开始讨论的一阶微分方程的一个特例, 但附带一个条件:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -2y, \quad y(0) = 5.$$

这意味着所求得的解不仅要满足微分方程, 还要保证当 $x = 0$ 时$y = 5$. 我们知道 $y = A\mathrm{e}^{kx}$ 是微分方程 $\mathrm{d}y/\mathrm{d}x = ky$ 的通解, 令 $k = -2$ 即知上面微分方程的通解是 $y = A\mathrm{e}^{-2x}$, A 为常数. 现在代入 $x = 0$ 和 $y = 5$ 可知 $5 = A\mathrm{e}^{-2(0)}$, 可得 $A = 5$. 附带的信息 $y(0) = 5$ 使得我们能够确定 A 的值, 所以实际解为 $y = 5\mathrm{e}^{-2x}$.

我们刚才看的是 IVP(Initial Value Problem, 初值问题) 的例子. 其思想是已知一个初始条件 (在这个例子中为 $y(0) = 5$) 和相关的微分方程 (在这个例子中为 $\mathrm{d}y/\mathrm{d}x = -2y$), 你可以运用这两个条件来求无不定常数的解. 对于一个二阶微分方程, 需积分两次, 所以你将得到两个不定常数, 由此可知需要两条已知信息. 一般地, 这两条信息是 $y(0)$ 的值和 $y'(0)$ 的值 ($x = 0$ 处的导数). 我们将在 30.4.2 节给出一些例题.

现在, 微分方程的研究已相当广泛. 这些问题很难求解, 事实上, 基本是不可能求解的, 至少一般是这样的. 幸运的是, 有一些简单的类型解起来并不很麻烦. 我们将讨论其中三种：一阶可分离变量方程、一阶线性方程、线性常系数方程.

30.2 可分离变量的一阶微分方程

如果能够把一阶微分方程中所有关于 y 的部分 (包括 $\mathrm{d}y$) 放在一边, 所有关于 x 的部分 (包括 $\mathrm{d}x$) 放在另一边, 则该微分方程被称为是**可分离变量的**. 例如, 方程 $\mathrm{d}y/\mathrm{d}x = ky$ 可重新整理为

$$\frac{1}{ky}\mathrm{d}y = \mathrm{d}x,$$

故它是可分离变量的. 另一个例子, 方程

$$\frac{\mathrm{d}y}{\mathrm{d}x} - \cos^2(y)\cos(x) = 0$$

可重新整理 (代数运算自行验证!) 成

$$\sec^2(y)\mathrm{d}y = \cos(x)\mathrm{d}x.$$



现在, 继续计算的方法是两边加积分号并积分, 然后整理[①]求 y. 在第一个例子中, 我们得到

$$\int \frac{1}{ky}\mathrm{d}y = \int \mathrm{d}x,$$

即

$$\frac{1}{k}\ln|y| = x + C,$$

其中 C 为常数. 要解 y, 两边乘 k 并取指数. 我们得到

$$|y| = \mathrm{e}^{kx+kC} = \mathrm{e}^{kC}\mathrm{e}^{kx}.$$

这意味着 $y = \pm\mathrm{e}^{kC}\mathrm{e}^{kx}$. 现在, $\pm\mathrm{e}^{kC}$ 是一些不为 0 的常数, 我们称它为 A, 因此给出了我们所期望的解 $y = A\mathrm{e}^{kx}$. (事实上, A 甚至可以为 0：若对所有 x 有 $y = 0$, 方程 $\mathrm{d}y/\mathrm{d}x = ky$ 显然可以成立, 因为两边都为 0. 这在我们的解中没有出现, 原因是我们要除以 y, 这是出于 y 恒不为 0 的假设.)

① 如你所预期的, 这些操作都可用链式求导法则证明.



对于第二个例子, 两边同时积分有

$$\int \sec^2(y)\mathrm{d}y = \int \cos(x)\mathrm{d}x,$$

可推出

$$\tan(y) = \sin(x) + C,$$

其中 C 是常数. 这个解已经很好了, 但或许你更愿意写成

$$y = \tan^{-1}(\sin(x) + C).$$

这里有个问题, 反正切函数的值域为 $(-\pi/2, \pi/2)$. 我们应该可以在上述表达式上加 π 的任何整数倍, 仍得到有效解. 事实上, $\sec^2(y)$ 有周期 π, 故全解应该为

$$y = \tan^{-1}(\sin(x) + C) + n\pi,$$

其中 C 为常数, n 为整数. 或许我们应该避开这些讨论, 仍写成 $\tan(y) = \sin(x) + C$ 的形式. (同样, 我们在开始求解时曾除以 $\cos^2(y)$, 这导致我们丢了常数解 $y = n\pi/2$, 其中 n 为奇数, 因为这些是当 $\cos^2(y) = 0$ 时的值. 这些解在上面解中 $C \to \pm\infty$ 时出现.)



对于同样的例子, 涉及初值时会是怎样的情形呢? 例如, 考虑 IVP

$$\frac{\mathrm{d}y}{\mathrm{d}x} - \cos^2(y)\cos(x) = 0, \quad y(0) = \frac{\pi}{4}.$$

若用上面的方法求解微分方程, 可得与前面一样的

$$\tan(y) = \sin(x) + C.$$

现在, 令 $x = 0$ 和 $y = \pi/4$ 可得

$$\tan(\pi/4) = \sin(0) + C,$$

这意味着 $C = 1$. 所以我们有

$$\tan(y) = \sin(x) + 1.$$

若我们写为

$$y = \tan^{-1}(\sin(x) + 1) + n\pi,$$

其中 n 为整数, 还令 $x = 0$ 和 $y = \pi/4$, 可知 $\pi/4 = \tan^{-1}(1) + n\pi$, 这意味着 $n = 0$. 故将解写成

$$y = \tan^{-1}(\sin(x) + 1)$$

是合情理的. 为使该点更清晰, 令初始条件为 $y(0) = 5\pi/4$ 而不是 $y(0) = \pi/4$. 将它代入方程 $\tan(y) = \sin(x) + C$, 又一次推出了 $C = 1$, 因为 $\tan(5\pi/4) = 1$. 所以, 我们再次求出了 $\tan(y) = \sin(x) + 1$, 但将这个方程写成 $y = \tan^{-1}(\sin(x) + 1)$ 是错误的. 为什么? 当 $x = 0$, 我们有

$$y = \tan^{-1}(\sin(0) + 1) = \tan^{-1}(0) = \frac{\pi}{4},$$

这不是我们想要的. 所以应加上 π, 即

$$y = \tan^{-1}(\sin(x) + 1) + \pi.$$

现在微分方程得到满足, 且如我们所希望的 $y(0) = 5\pi/4$. 如果初始条件为 $y(0) = \pi/4 + n\pi$ 对任意非 0 整数 n 成立, 同样需要警惕. 这需要微妙技巧!

30.3 一阶线性方程

这是另一种一阶微分方程：

$$\frac{dy}{dx} + p(x)y = q(x),$$

其中 p 和 q 是关于 x 的函数. 这样的方程称为**一阶线性微分方程**. 它可能不是可分离变量的, 甚至连线性看起来也不很明显! 例如,

$$\frac{dy}{dx} + 6x^2y = e^{-2x^3}\sin(x)$$

就不像是线性的, 然而这个方程确实是一阶线性的, 因为 y 和 dy/dx 的幂次都是 1. 而方程

$$\frac{dy}{dx} + 6x^2y^3 = e^{-2x^3}\sin(x)$$

不是一阶线性的, 因为 y^3 不是 y 的一次. 类似地,

$$\left(\frac{dy}{dx}\right)^2 + 6x^2y = e^{-2x^3}\sin(x)$$

也不是线性的, 因为量 dy/dx 被平方了.



回到前面的线性方程

$$\frac{dy}{dx} + 6x^2y = e^{-2x^3}\sin(x).$$

这个方程不是可分离变量的. 试一下! 你不可能得到一边都关于 y 而另一边都关于 x 的方程. 幸运的是, 有个诀窍可用. 假如我们两边都乘 e^{2x^3}. 这个操作显然会使右边变得简洁了, 但其实还有一个更有趣的作用. 我们来看看发生了什么：

$$e^{2x^3}\frac{dy}{dx} + 6x^2e^{2x^3}y = \sin(x).$$

现在要仔细了：我将它改写为了

$$\frac{d}{dx}(e^{2x^3}y) = \sin(x),$$

这其实没隐藏什么. 怎么可能呢? 好吧, 我所做的就是在求导时把乘积法则反过来用了一下罢了! (小菜一碟.) 为了证明这是正确的, 你要做的就是把导数算出来. 事实上, 根据乘积法则, 其中一项为 e^{2x^3} 乘以 y 的导数, 即 $y \times 6x^2e^{2x^3}$ (用链式求导法则). 那正是原来方程的左边! 所以我们确实有

$$\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{2x^3}y) = \sin(x).$$

现在要做的就是两边关于 x 积分. 这样, 就消掉了左边的导数, 剩下

$$\mathrm{e}^{2x^3}y = \int \sin(x)\mathrm{d}x = -\cos(x) + C.$$

除以 e^{2x^3}, 得到解



$$y = (C - \cos(x))\mathrm{e}^{-2x^3},$$

其中 C 为任意常数. 现在可以试着对其求导来验证它满足原微分方程!

上述求解的关键是乘以 e^{2x^3}. 之后, 我们能将左边整体写成 $\dfrac{\mathrm{d}}{\mathrm{d}x}$ (某式) 的形式, 这样就很容易积分了. 出于这个原因, e^{2x^3} 被称为**积分因子**. 可知, 对于一般一阶线性微分方程

$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = q(x),$$

一个好的积分因子由等式

$$\text{积分因子} = \mathrm{e}^{\int p(x)\mathrm{d}x}$$

给出, 这里不必对积分结果 $+C$. 在将原微分方程乘了这个积分因子之后, 左边就可被 “因式分解” 成

$$\frac{\mathrm{d}}{\mathrm{d}x}(\text{积分因子} \times y).$$

我们稍后讨论原因. 现在, 我们用这个更一般的框架来重新计算前面的例子

$$\frac{\mathrm{d}y}{\mathrm{d}x} + 6x^2y = \mathrm{e}^{-2x^3}\sin(x)$$

首先, 通过取 y 的系数 (即 $6x^2$) 来找积分因子, 对其积分并指数化结果：

$$\text{积分因子} = \mathrm{e}^{\int 6x^2\mathrm{d}x} = \mathrm{e}^{2x^3}.$$

现在可以如前那样：用 e^{2x^3} 乘微分方程并将左边重写为 $\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{2x^3}y)$, 它是积分因子和 y 乘积的导数.



目前为止, 学习这个方法的最好办法就是做大量的练习, 直到掌握为止. 这里有另外两个例子. 首先, 如何解

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{x}y + \mathrm{e}^{2x}, \quad y(0) = 2(\mathrm{e} - 1)?$$

这是一个 IVP, 但我们担心的是微分方程解完后的事. 第一件事是将其变成标准形式, 意思是需将所有关于 y 的部分放在左边, 所有关于 x 的部分放在右边, 且 $\mathrm{d}y/\mathrm{d}x$ 的系数要为 1. 在本例中, 我们只需两边减去 $\mathrm{e}^{x}y$, 得

$$\frac{\mathrm{d}y}{\mathrm{d}x} - \mathrm{e}^{x}y = \mathrm{e}^{2x}, \quad y(0) = 2(\mathrm{e} - 1).$$

y 的系数为 $-\mathrm{e}^x$, 故积分因子是该量积分的指数化：

$$\text{积分因子} = \mathrm{e}^{\int(-\mathrm{e}^x)\mathrm{d}x} = \mathrm{e}^{-\mathrm{e}^x}.$$

(记住, 这里不需 $+C$.) 我们用这个积分因子乘上述微分方程的两边：

$$\mathrm{e}^{-\mathrm{e}^x}\frac{\mathrm{d}y}{\mathrm{d}x} - \mathrm{e}^x\mathrm{e}^{-\mathrm{e}^x}y = \mathrm{e}^{-\mathrm{e}^x}\mathrm{e}^{2x}.$$

如往常一样, 左边是 y 乘积分因子的导数, 所以我们有

$$\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{-\mathrm{e}^x}y) = \mathrm{e}^{-\mathrm{e}^x}\mathrm{e}^{2x}.$$

最好通过对左边求导来验证这个化简的合理性. 总之, 对上述方程的两边积分可得

$$\mathrm{e}^{-\mathrm{e}^x}y = \int \mathrm{e}^{-\mathrm{e}^x}\mathrm{e}^{2x}\mathrm{d}x,$$



为了求这个积分, 令 $t = \mathrm{e}^x$, 则 $\mathrm{d}t = \mathrm{e}^x\mathrm{d}x$. 注意, 需将 e^{2x} 写成 $\mathrm{e}^x\mathrm{e}^x$ 来计算. 我把积分的计算 (运用分部积分法) 留给你来完成, 并验证结果方程为

$$\mathrm{e}^{-\mathrm{e}^x}y = -\mathrm{e}^x\mathrm{e}^{-\mathrm{e}^x} - \mathrm{e}^{-\mathrm{e}^x} + C.$$

最后, 两边除以积分因子 $\mathrm{e}^{-\mathrm{e}^x}$ 可得

$$y = -\mathrm{e}^x - 1 + C\mathrm{e}^{\mathrm{e}^x},$$

对某常数 C 成立. 现在剩下的就是解 IVP. 当 $x = 0$ 时, 我们知道 $y = 2(\mathrm{e}-1)$, 所以将这个代入上述方程, 有

$$2(\mathrm{e}-1) = -\mathrm{e}^0 - 1 + C\mathrm{e}^{\mathrm{e}^0}.$$

你可以很容易解出 $C = 2$, 故最后的解是

$$y = 2\mathrm{e}^{\mathrm{e}^x} - \mathrm{e}^x - 1.$$

可对其求导来验证它满足原微分方程.



我们再快速浏览一个一阶线性微分方程

$$\tan(x)\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{\sin(x)} - y.$$

首先, 将关于 y 的部分放在左边并令方程除以 $\tan(x)$, 以使 $\mathrm{d}y/\mathrm{d}x$ 的系数等于 1：

$$\frac{\mathrm{d}y}{\mathrm{d}x} + \cot(x)y = \cot(x)\mathrm{e}^{\sin(x)}.$$

y 的系数是 $\cot(x)$, 故

$$\text{积分因子} = \mathrm{e}^{\int\cot(x)\mathrm{d}x} = \mathrm{e}^{\ln(\sin(x))} = \sin(x).$$

(技术上讲, 我们应该写成 $|\sin(x)|$, 但这会使事情变得不必要的复杂.) 不管怎么说, 用 $\sin(x)$ 乘微分方程可得

$$\sin(x)\frac{\mathrm{d}y}{\mathrm{d}x} + \cos(x)y = \cos(x)\mathrm{e}^{\sin(x)},$$

因为 $\sin(x)\cot(x) = \cos(x)$. 现在左边变为 y 乘积分因子的导数 (验证它)：

$$\frac{\mathrm{d}}{\mathrm{d}x}(y\sin(x)) = \cos(x)\mathrm{e}^{\sin(x)}.$$

两边积分 (用换元来化简右边)：

$$y\sin(x) = \int \cos(x)\mathrm{e}^{\sin(x)}\mathrm{d}x = \mathrm{e}^{\sin(x)} + C.$$

最后, 用 $\sin(x)$ 除以两边可得

$$y = \csc(x)\mathrm{e}^{\sin(x)} + C\csc(x),$$



我们已经找到了微分方程的解.

综述, 下面是求解一阶线性微分方程的方法.

- 将包含 y 的部分放在左边, 包含 x 的部分放在右边, 然后两边除以 $\mathrm{d}y/\mathrm{d}x$ 的系数得到一个标准形式的方程
$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = q(x).$$
- 两边乘积分因子, 我们称其为 $f(x)$, 它由
$$\boxed{\text{积分因子 } f(x) = \mathrm{e}^{\int p(x)\mathrm{d}x}}$$
给出, 这里不需为指数上的积分 $+C$.
- 左边变为 $\dfrac{\mathrm{d}}{\mathrm{d}x}(f(x)y)$, 其中 $f(x)$ 是积分因子. 用这个新的左边重写方程.
- 两边积分, 这次必须在右边 $+C$.
- 除以积分因子来解出 y.

练习这个方法, 你不会后悔的!



为什么积分因子起作用

为什么怪异的表达式 $\mathrm{e}^{\int p(x)\mathrm{d}x}$ 是个很好的积分因子呢? 假设我们取一般方程

$$\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = q(x),$$

并用积分因子 $\mathrm{e}^{\int p(x)\mathrm{d}x}$ 乘以它. 我们得到

$$\mathrm{e}^{\int p(x)\mathrm{d}x}\frac{\mathrm{d}y}{\mathrm{d}x} + \mathrm{e}^{\int p(x)\mathrm{d}x}p(x)y = \text{关于 } x \text{ 的部分}.$$

现在我只关注左边, 故只将右边写为：关于 x 的部分. 我们已经声明可将左边重写, 使得上面的方程变为

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{\int p(x)\mathrm{d}x}y\right) = \text{关于 } x \text{ 的部分},$$

这就容易求解了. 为了证明我们的声明, 对左边运用乘积法则, 写为

$$\mathrm{e}^{\int p(x)\mathrm{d}x}\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{\int p(x)\mathrm{d}x}\right)y.$$

这个几乎就是我们需要的了, 我们只需运用链式求导法则写成

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{\int p(x)\mathrm{d}x}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\int p(x)\mathrm{d}x\right) \times \mathrm{e}^{\int p(x)\mathrm{d}x} = p(x)\mathrm{e}^{\int p(x)\mathrm{d}x}.$$

注意, $\frac{\mathrm{d}}{\mathrm{d}x}\int p(x)\mathrm{d}x = p(x)$, 因为 $\int p(x)\mathrm{d}x$ (不带 $+C$) 是 p 的反导. 现在如果把前面各部分整理一下, 最后可知

$$\mathrm{e}^{\int p(x)\mathrm{d}x}\frac{\mathrm{d}y}{\mathrm{d}x} + \mathrm{e}^{\int p(x)\mathrm{d}x}p(x)y = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{\int p(x)\mathrm{d}x}y\right).$$

我们的方法是可行的!

30.4 常系数微分方程

现在我们来讨论常系数线性微分方程. 这些方程形如

$$a_n\frac{\mathrm{d}^n y}{\mathrm{d}x^n} + \cdots + a_2\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + a_1\frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = f(x),$$

这里 f 是只关于 x 的函数, $a_n, \cdots, a_1, a_0$ 只是一些普通的常实数. 注意上面方程的左边有点像一个 y 的多项式, 不过它用的是导数而不是 y 的幂.

我们来看个例题. 考虑微分方程

$$3\frac{\mathrm{d}y}{\mathrm{d}x} - \sin(5x) = 12x - 6y.$$

这个方程可以整理成所有关于 x 的部分在右边, 所有关于 y 的部分 (包括导数) 在左边的形式. 最后, 除以 3 可得

$$\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = 4x + \frac{1}{3}\sin(5x).$$

这是一个一阶常系数线性方程. 其实, 你可以用前一节的一阶线性方程的方法来解它. 如果这么做, 你将需要用到积分因子, 而这在这个例子中有点难度 (试一下看看). 我们很快会讨论求解该方程的另一个方法. 事实上, 我们将在 30.4.6 节求解上面这个例子.

我们也要详细考察一下二阶的情形. 在这种情况下, 我们要讨论形如

$$a\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = f(x)$$

的方程. 例如,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - 5\frac{\mathrm{d}y}{\mathrm{d}x} + 6y = 2x^2\mathrm{e}^x.$$

我们将在 30.4.6 节讨论它的解法. 首先, 我们需要讨论一阶和二阶常系数线性方程的一般解法[1].

我们从一个简单例子入手：假设右边没有关于 x 的部分, 例如

① 这些方法对高阶方程也适用, 不过本书将主要着重于一阶和二阶方程.

$$\frac{\mathrm{d}y}{\mathrm{d}x} - 3y = 0 \quad 和 \quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2} - \frac{\mathrm{d}y}{\mathrm{d}x} + 20y = 0.$$

这样的方程称为齐次的. 我们来看如何求解一阶 (左边的例子) 和二阶 (右边的例子) 齐次方程.

30.4.1 解一阶齐次方程

这个非常简单.

$$\frac{\mathrm{d}y}{\mathrm{d}x} + ay = 0$$

的解为 $y = A\mathrm{e}^{-ax}$. (其实, 这个方程就是 $\mathrm{d}y/\mathrm{d}x = ky$, 其中 $k = -a$ 的形式, 见 30.1 节和 30.2 节.) 例如, 给定微分方程

$$\frac{\mathrm{d}y}{\mathrm{d}x} - 3y = 0,$$

可以直接写出其解为 $y = A\mathrm{e}^{3x}$, 其中 A 是常数.

30.4.2 解二阶齐次方程

这种情况有点棘手. 我们需解

$$a\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = 0.$$

它看起来有点奇怪, 最简单的办法就是提取出一个二次方程. 这个二次方程称为*特征二次方程*, 即 $at^2 + bt + c = 0$. 例如, 考虑下面 3 个微分方程:

(a) $y'' - y' - 20y = 0$; (b) $y'' + 6y' + 9y = 0$; (c) $y'' - 2y' + 5y = 0$;

注意, 我们已经用 y' 代替 $\mathrm{d}y/\mathrm{d}x$, 用 y'' 代替 $\mathrm{d}^2y/\mathrm{d}x^2$. 不管怎样, 这 3 个例子的特征方程分别为 $t^2 - t - 20 = 0$、$t^2 + 6t + 9 = 0$ 和 $t^2 - 2t + 5 = 0$.

接下来就是求特征方程的根. 这有三种可能, 取决于方程是否有两个实根、一个 (双重) 实根或两个复根. 我们来总结一下整个方法, 然后解上述三个例子.

下面是解齐次方程 $ay'' + by' + cy = 0$ 的方法.

(1) 写出特征二次方程 $at^2 + bt + c = 0$ 并解 t.

(2) 若有两个不同实根 α 和 β, 解为

$$y = A\mathrm{e}^{\alpha x} + B\mathrm{e}^{\beta x}.$$

(3) 若只有一个 (双重) 实根 α, 解为

$$y = A\mathrm{e}^{\alpha x} + Bx\mathrm{e}^{\alpha x}.$$

(4) 若有两个复根, 它们将是共轭的, 即其形为 $\alpha \pm i\beta$. 解为

$$y = \mathrm{e}^{\alpha x}(A\cos(\beta x) + B\sin(\beta x)).$$

在所有情形中 (2、3 和 4), A 和 B 为不定常数.

所以, 对前面的例 (a), 我们看到特征二次方程是 $t^2 - t - 20 = 0$. 若将二次式因式分解为 $(t + 4)(t - 5)$, 显然方程的解为 $t = -4$ 和 $t = 5$. 由上面第 (2) 步可知,

方程 $y''-y'-20y=0$ 的解为

$$y=A\mathrm{e}^{-4x}+B\mathrm{e}^{5x},$$

对某些常数 A 和 B 成立.

例 (b) 的特征二次方程 $t^2+6t+9=0$ 化简为 $(t+3)^2=0$, 因此唯一解为 $t=-3$. 由前面第 (3) 步, 齐次方程 $y''+6y'+9=0$ 的解为

$$y=A\mathrm{e}^{-3x}+Bx\mathrm{e}^{-3x}.$$



最后, 如果我们用二次公式来解例 (c) 的特征二次方程 $t^2-2t+5=0$, 可得 $t=1\pm 2\mathrm{i}$. (试一下看看!) 故, 由 $\alpha=1$ 和 $\beta=2$, 上面的第 (4) 步给出了 $y''-2y'+5y=0$ 的解:

$$y=\mathrm{e}^{x}(A\cos(2x)+B\sin(2x)).$$

同样, A 和 B 为不定常数.

30.4.3 为什么特征二次方程适用



为什么前面的方法适用呢? (若你不关心原因, 可以直接转到下一节!) 考虑将 $y=\mathrm{e}^{\alpha x}$ 代入方程 $ay''+by'+cy=0$ 时会发生什么. 我们有 $y'=\alpha\mathrm{e}^{\alpha x}$ 和 $y''=\alpha^2\mathrm{e}^{\alpha x}$, 所以



$$ay''+by'+cy=a\alpha^2\mathrm{e}^{\alpha x}+b\alpha\mathrm{e}^{\alpha x}+c\mathrm{e}^{\alpha x}=(a\alpha^2+b\alpha+c)\mathrm{e}^{\alpha x}.$$

故, 若 α 为特征二次式 at^2+bt+c 的一个根, 则有 $a\alpha^2+b\alpha+c=0$. 上述等式暗示了 $ay''+by'+cy=0$, 即 $y=\mathrm{e}^{\alpha x}$ 解出了微分方程! 同样, 该解的任何常数倍也是方程的解, 且若有另一个根 β, 则可将两个解 $y=A\mathrm{e}^{\alpha x}$ 和 $y=B\mathrm{e}^{\beta x}$ 加起来得到更多解. (试试看!) 但要小心第 (2) 步.

下面我们来看第 (4) 步. 若二次方程的两个解是形如 $\alpha+\mathrm{i}\beta$ 的共轭复根, 则根据第 (2) 步的讨论, 解定为

$$y=A\mathrm{e}^{(\alpha+\mathrm{i}\beta)x}+B\mathrm{e}^{(\alpha-\mathrm{i}\beta)x}=\mathrm{e}^{\alpha x}(A\mathrm{e}^{\mathrm{i}\beta x}+B\mathrm{e}^{-\mathrm{i}\beta x}),$$

这里 A 和 B 甚至可以为复数. 现在可用欧拉等式 (见 28.2 节) 得

$$\begin{aligned}y&=\mathrm{e}^{\alpha x}(A(\cos(\beta x)+\mathrm{i}\sin(\beta x))+B(\cos(\beta x)-\mathrm{i}\sin(\beta x)))\\&=\mathrm{e}^{\alpha x}((A+B)\cos(\beta x)+(A-B)\mathrm{i}\sin(\beta x)).\end{aligned}$$

重新标记常数 $(A+B)$ 为 A, 常数 $(A-B)\mathrm{i}$ 为 B, 得到正确的公式.

最后对第 (3) 步, 假定特征二次方程只有一个根 α. 若将 $y=x\mathrm{e}^{\alpha x}$ 代入微分方程 $ay''+by'+cy=0$, 可以由 $y'=\alpha x\mathrm{e}^{\alpha x}+\mathrm{e}^{\alpha x}$ 和 $y''=\alpha^2x\mathrm{e}^{\alpha x}+2\alpha\mathrm{e}^{\alpha x}$ 推出

$$ay''+by'+cy=(a\alpha^2+b\alpha+c)x\mathrm{e}^{\alpha x}+(2a\alpha+b)\mathrm{e}^{\alpha x}.$$

若 α 是 at^2+bt+c 的双重根, 则不仅 $a\alpha^2+b\alpha+c=0$, 而且 $2a\alpha+b=0$[①]. 由此可

---

① 这是二次方程 $at^2+bt+c=0$ 有双重根 $t=\alpha$ 时 $2a\alpha+b=0$ 的原因: 判别式为 0, 所以 $b^2=4ac$. 则

$$(2a\alpha+b)^2=4a^2\alpha^2+4ab\alpha+b^2=4a^2\alpha^2+4ab\alpha+4ac=4a(a\alpha^2+b\alpha+c)=0.$$

又因为 $(2a\alpha+b)^2=0$, 当然有 $2a\alpha+b=0$.

推出前面第 (3) 步的正确解.

30.4.4 非齐次方程和特解

我们来看方程右边仅有 x 部分时的情况. 例如, 考虑微分方程

$$y''-y'-20y=\mathrm{e}^x$$

它不是齐次的, 因为右边有 e^x. 试着猜一个解, 我们知道 e^x 的所有导数为 e^x, 试着令 $y=\mathrm{e}^x$. 则 $y'=\mathrm{e}^x$, $y''=\mathrm{e}^x$, 所以左边 $y''-y'-20y$ 变为 $\mathrm{e}^x-\mathrm{e}^x-20\mathrm{e}^x=-20\mathrm{e}^x$, 不等于右边, 但很接近. 我们只需除以 -20. 再试一次：令 $y=-\dfrac{1}{20}\mathrm{e}^x$. 则 y' 和 y'' 也为 $-\dfrac{1}{20}\mathrm{e}^x$, 所以我们有

$$y''-y'-20y=-\frac{1}{20}\mathrm{e}^x-\left(-\frac{1}{20}\mathrm{e}^x\right)-20\left(-\frac{1}{20}\mathrm{e}^x\right)=\mathrm{e}^x.$$

我们证明了 $y=-\dfrac{1}{20}\mathrm{e}^x$ 是原方程 $y''-y'-20y=\mathrm{e}^x$ 的一个解, 但它不是唯一解. 要知道原因, 考虑相关齐次方程

$$y''-y'-20y=0.$$

这其实是 30.4.2 节的例 (a). 我们知道全解为

$$y=A\mathrm{e}^{-4x}+B\mathrm{e}^{5x}.$$

因此我们来做个小游戏. 我们将用 y_H 代替 y 来写这个解, 其中 H 表示齐次. 我们已经证明了

$$\text{若 } y_H=A\mathrm{e}^{-4x}+B\mathrm{e}^{5x},\quad \text{则} y_H''-y_H'-20y_H=0.$$

另一方面, 我们在前面说明了

$$\text{若 } y_P=-\frac{1}{20}\mathrm{e}^x,\quad \text{则 } y_P''-y_P'-20y_P=\mathrm{e}^x.$$

这里我把前面的解 $-\dfrac{1}{20}\mathrm{e}^x$ 写为 y_P, 称其为特解, 它解释了下标为何是 P. 现在, 如果把方程 $y_H''-y_H'-20y_H=0$ 和 $y_P''-y_P'-20y_P=\mathrm{e}^x$ 加起来, 把导数放在一起, 我们得到

$$y_H''+y_P''-y_H'-y_P'-20y_H-20y_P=0+\mathrm{e}^x.$$

事实上, 由于导数之和等于和的导数, 对二阶导也一样, 我们可得

$$(y_H+y_P)''-(y_H+y_P)'-20(y_H+y_P)=\mathrm{e}^x.$$

因此, 若 $y=y_H+y_P$, 则 y 也是原微分方程 $y''-y'-20y=\mathrm{e}^x$ 的一个解. 换句话说, 我们可以取特解

$$y_P=-\frac{1}{20}\mathrm{e}^x,$$

它确实是原微分方程的解, 然后加上微分方程齐次形式的任意解, 结果仍为原微分方程的解. 进一步地, 非齐次方程的所有解均为该形式.

一阶和二阶微分方程都可用这个方法. 唯一的问题是怎么猜这个特解. 在下一节, 我们将讨论如何推测解的形式 (与 18.3 节中的部分分式法类似). 若幸运的话, 可以代入该形式并求出未知常数来确定特解.



下面我们总结一下所讨论的方法.

(1) 将方程整理成正确的形式, 即将所有含 x 的部分放在右边. 则可将一阶形式方程化简为

$$\frac{\mathrm{d}y}{\mathrm{d}x}+ay=f(x)$$

或二阶形式化简为

$$a\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+b\frac{\mathrm{d}y}{\mathrm{d}x}+cy=f(x).$$

(2) 运用 30.4.1 节和 30.4.2 节的方法, 解相应的齐次方程

$$\frac{\mathrm{d}y}{\mathrm{d}x}+ay=0 \quad 或 \quad a\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+b\frac{\mathrm{d}y}{\mathrm{d}x}+cy=0.$$

我们将解记作 y_H, 它有一个或两个待定常数 (取决于方程是一阶还是二阶). 我们称 y_H 为方程的齐次解.

(3) 若原函数 f 为 0, 则计算结束, 全解为 $y=y_H$.

(4) 另一方面, 若函数 f 不为 0, 则写出特解 y_P 的形式 (见 30.4.5 节). 这个形式有一些需要确定的常数. 将 y_P 代入原方程并令系数相等来求待定常数.

(5) 最后, 解为 $y=y_H+y_P$.

我们将在 30.4.8 节讨论 IVP 的情况. 现在, 我们来看如何求特解.

30.4.5 求特解



目前为止, 我们忽略了可能出现在右边的含 x 部分 (之前称为 $f(x)$), 现在该讨论这部分了. 其方法是写出特解的形式, 然后将该形式代入方程来求真正的解. 通过后面的表格可知如何写出正确的形式. 例如, 在微分方程

$$y'-3y=5\mathrm{e}^{2x}$$

中, 右边是 e^{2x} 的倍数, 由表可知特解的形式应为 $y_P=C\mathrm{e}^{2x}$, 其中 C 是一个常数, 我们需将 y_P 代入原方程来求出这个常数. 易知 $y_P'=2C\mathrm{e}^{2x}$, 因此有

$$2C\mathrm{e}^{2x}-3(C\mathrm{e}^{2x})=5\mathrm{e}^{2x}.$$

它可化简为 $-C\mathrm{e}^{2x}=5\mathrm{e}^{2x}$, 所以 $C=-5$. 由此, 特解为 $y_P=-5\mathrm{e}^{2x}$. 事实上, 由于我们在 30.4.1 节见过齐次形式 $y'-3y=0$ 的解为 $y_H=A\mathrm{e}^{3x}$, 因而可以知道 $y'-3y=5\mathrm{e}^{2x}$ 的全解是

$$y=y_H+y_P=A\mathrm{e}^{3x}-5\mathrm{e}^{2x},$$

其中 A 为未知常数. 注意, 齐次解包含未知常数, 而特解一定不能含未知常数.

下面就是那个表格.

若 f 是一个	则形式为
次数为 n 的多项式	y_P =**次数为 n 的一般多项式**
例, $f(x)=7$	$y_P=a$
$f(x)=3x-2$	$y_P=ax+b$
$f(x)=10x^2$	$y_P=ax^2+bx+c$
$f(x)=-x^3-x^2+x+22$	$y_P=ax^3+bx^2+cx+d$
指数 e^{kx} 的倍数	$y_P=C\mathrm{e}^{kx}$
例, $f(x)=10\mathrm{e}^{-4x}$	$y_P=C\mathrm{e}^{-4x}$
$f(x)=\mathrm{e}^x$	$y_P=C\mathrm{e}^x$
$\cos(kx)$ 的倍数+$\sin(kx)$ 的倍数	$y_P=C\cos(kx)+D\sin(kx)$
例, $f(x)=2\sin(3x)-5\cos(3x)$	$y_P=C\cos(3x)+D\sin(3x)$
$f(x)=\cos(x)$	$y_P=C\cos(x)+D\sin(x)$
$f(x)=2\sin(11x)$	$y_P=C\cos(11x)+D\sin(11x)$
上面某些形式的和或积	**这些形式的和或积 (若为积, 删掉一个常数)**
例, $f(x)=2x^2+\mathrm{e}^{-6x}$	$y_P=ax^2+bx+c+C\mathrm{e}^{-6x}$
$f(x)=2x^2\mathrm{e}^{-6x}$	$y_P=(ax^2+bx+c)\mathrm{e}^{-6x}$
$f(x)=7\mathrm{e}^{2x}\sin(3x)$	$y_P=(C\cos(3x)+D\sin(3x))\mathrm{e}^{2x}$
$f(x)=\cos(2x)+6\sin(x)$	$y_P=C\cos(2x)+D\sin(2x)+E\cos(x)+F\sin(x)$
$f(x)=4x\cos(3x)$	$y_P=(x+b)(C\cos(3x)+D\sin(3x))$
若 y_P 与 y_H 冲突, 令特解的形式乘以 x 或 x^2	

除了最后一行, 这个表对于 “若为积, 删掉一个常数” 是不言自明的. 最后一行将在 30.4.7 节加以解释. 要明白这个不言自明, 首先注意到两种形式乘起来时有一个多余的常数. 例如, $2x^2\mathrm{e}^{-6x}$ 看似会引入形式 $(ax^2+bx+c)C\mathrm{e}^{-6x}$, 但常数 C 是没必要的, 可以删除, 因为它可以并入到其他的常数 a、b 和 c 中. 这点同样适用于表中的例子 $7\mathrm{e}^{2x}\sin(3x)$ 和 $4x\cos(3x)$.

(顺便说一下, 这个表只显示了若 f 为多项式、指数、正弦、余弦, 或一个或多个这些类型函数的积或和的方法. 这个方法不适用于其他情形. 还有一个更一般的方法 “参数变异法”, 但它不在本书的讨论范围内.)

30.4.6 求特解的例子

写出了 y_P 的形式后, 还需将其代入原微分方程来求常数. 为使计算更容易, 首先要求 y_P' 和 y_P''(对一阶情形, 只需求 y_P'). 我们来看一个这样的例子, 然后再返回完成 30.4 节中未完成的两个例子.



首先考虑微分方程

$$y''-4y'+4y=25\mathrm{e}^{3x}\sin(2x).$$

我们来快速搞定齐次部分. 其实, $y''-4y'+4y=0$ 的特征二次方程 $t^2-4t+4=0$ 只有一个解, 即 $t=2$. 因此, 我们有 $y_H=A\mathrm{e}^{2x}+Bx\mathrm{e}^{2x}$, 其中 A 和 B 为常数. 现在我们来找特解. 将微分方程右边的 $25\mathrm{e}^{3x}\sin(2x)$ 分成两部分：$25\mathrm{e}^{3x}$ 和 $\sin(2x)$. 根据前面的表, e^{3x} 的常数倍形式为 $C\mathrm{e}^{3x}$, $\sin(2x)$ 的形式为 $C\cos(2x)+D\sin(2x)$. 我

们要将这些乘在一起, 不过在这个过程中可将常数合并写成

$$y_P = \mathrm{e}^{3x}(C\cos(2x) + D\sin(2x)).$$

现在多次运用乘积法则来做一些烦琐的计算:

$$\begin{aligned}
y_P &= \mathrm{e}^{3x}(C\cos(2x) + D\sin(2x)),\\
y_P' &= \mathrm{e}^{3x}(-2C\sin(2x) + 2D\cos(2x)) + 3\mathrm{e}^{3x}(C\cos(2x) + D\sin(2x))\\
&= \mathrm{e}^{3x}((3C+2D)\cos(2x) + (3D-2C)\sin(2x)),\\
y_P'' &= \mathrm{e}^{3x}(-2(3C+2D)\sin(2x) + 2(3D-2C)\cos(2x))\\
&\qquad + 3\mathrm{e}^{3x}((3C+2D)\cos(2x) + (3D-2C)\sin(2x))\\
&= \mathrm{e}^{3x}((5C+12D)\cos(2x) + (5D-12C)\sin(2x)).
\end{aligned}$$

现在该将这些代入原微分方程 $y'' - 4y' + 4y = 25\mathrm{e}^{3x}\sin(2x)$ 了. 我们得到了看起来很长的方程

$$\begin{aligned}
\mathrm{e}^{3x}((5C+12D)\cos(2x) &+ (5D-12C)\sin(2x))\\
&- 4\mathrm{e}^{3x}((3C+2D)\cos(2x) + (3D-2C)\sin(2x))\\
&+ 4\mathrm{e}^{3x}(C\cos(2x) + D\sin(2x)) = 25\mathrm{e}^{3x}\sin(2x),
\end{aligned}$$

它可化简为

$$\mathrm{e}^{3x}(4D-3C)\cos(2x) + \mathrm{e}^{3x}(-4C-3D)\sin(2x) = 25\mathrm{e}^{3x}\sin(2x).$$

为了使这个表达式对所有 x 成立, $\mathrm{e}^{3x}\cos(2x)$ 部分需为 0 且 $\mathrm{e}^{3x}\sin(2x)$ 的系数需为 25. 这意味着 $4D-3C=0$ 且 $-4C-3D=25$. 同时解这些方程, 可得 $C=-4$ 和 $D=-3$. 现在我们知道 $y_P = \mathrm{e}^{3x}(-4\cos(2x) - 3\sin(2x))$, 故全解为

$$y = y_H + y_P = A\mathrm{e}^{2x} + Bx\mathrm{e}^{2x} - \mathrm{e}^{3x}(4\cos(2x) + 3\sin(2x)),$$

其中 A 和 B 为常数.

现在该遵照承诺完成 30.4 节的两个例子了:

$$y' + 2y = 4x + \frac{1}{3}\sin(5x) \quad 和 \quad y'' - 5y' + 6y = 2x^2\mathrm{e}^x.$$

你应该先试着解这两个方程. 如果完成了, 继续往下读.



左边的例子是一个一阶方程. 齐次形式为 $y' + 2y = 0$, 有解 $y = A\mathrm{e}^{-2x}$, 其中 A 为常数. 根据前面的表, 我们知道其特解形式为 $y_P = ax + b + C\cos(5x) + D\sin(5x)$. 我们需知导数, 即 $y_P' = a - 5C\sin(5x) + 5D\cos(5x)$. 将 y_P' 和 y_P 代入原方程, 可得

$$(a - 5C\sin(5x) + 5D\cos(5x)) + 2(ax + b + C\cos(5x) + D\sin(5x)) = 4x + \frac{1}{3}\sin(5x),$$

它可化简为

$$2ax + 2b + a + (5D+2C)\cos(5x) + (2D-5C)\sin(5x) = 4x + \frac{1}{3}\sin(5x).$$

现在要令该表达式中各部分的系数相等. 左边 x 的系数是 $2a$, 右边为 4, 故 $a=2$. 左边的常数为 $2b+a$, 而右边没有常数, 故 $2b+a=0$. 这就意味着 $b=-1$. 同时,

右边没有关于 $\cos(5x)$ 的项, 故 $5D + 2C = 0$. 还有, $\sin(5x)$ 项也必须对应, 故 $2D - 5C = 1/3$. 同时解最后这两个方程 (试试!) 可得 $C = -5/87$ 和 $D = 2/87$. 因此, 我们有

$$y_P = 2x - 1 - \frac{5}{87}\cos(5x) + \frac{2}{87}\sin(5x).$$

把这些整理后, 得到解

$$y = y_H + y_P = Ae^{-2x} + 2x - 1 - \frac{5}{87}\cos(5x) + \frac{2}{87}\sin(5x),$$

其中 A 为常数.

那另一个例子呢? 那是一个二阶方程, 其齐次形式为 $y'' - 5y' + 6y = 0$. 特征二次方程是 $t^2 - 5t + 6 = 0$, 其解为 $t = 2$ 和 $t = 3$. 因此, $y_H = Ae^{2x} + Be^{3x}$, 其中 A 和 B 为常数. 现在该求特解了. 由于原微分方程的右边为 $2x^2e^x$, 特解形式应该为 $y_P = (ax^2 + bx + c)e^x$, 要知道 e^x 的外面是不需加常数的, 因为那个常数可并入 a、b 和 c 中. 我们对 y_P 求两次导, 得到

$$\begin{aligned} y_P &= (ax^2 + bx + c)e^x, \\ y_P' &= (ax^2 + bx + c)e^x + (2ax + b)e^x \\ &= (ax^2 + (2a + b)x + (b + c))e^x, \\ y_P'' &= (ax^2 + (2a + b)x + (b + c))e^x + (2ax + (2a + b))e^x \\ &= (ax^2 + (4a + b)x + (2a + 2b + c))e^x. \end{aligned}$$

将其代入原方程 $y'' - 5y' + 6y = 2x^2e^x$ 可得

$$(ax^2 + (4a + b)x + (2a + 2b + c))e^x - 5(ax^2 + (2a + b)x + (b + c))e^x + 6(ax^2 + bx + c)e^x = 2x^2e^x.$$

它可化简成

$$(2ax^2 + (-6a + 2b)x + (2a - 3b + 2c))e^x = 2x^2e^x.$$

令系数相等可知, $2a = 2, -6a + 2b = 0$, $2a - 3b + 2c = 0$. 由此求得 $a = 1$, $b = 3$, $c = \frac{7}{2}$, 因此 $y_P = \left(x^2 + 3x + \frac{7}{2}\right)e^x$. 整个方程的解为

$$y = y_H + y_P = Ae^{2x} + Be^{3x} + \left(x^2 + 3x + \frac{7}{2}\right)e^x,$$

其中 A 和 B 为常数.

30.4.7　解决 y_P 和 y_H 间的冲突

30.4.5 节中表的最后一行指出, y_P 和 y_H 可能会有冲突. 为什么会这样呢? 考虑微分方程

$$y'' - 3y' + 2y = 7e^{2x}.$$

它的齐次形式为 $y'' - 3y' + 2y = 0$, 特征二次方程为 $t^2 - 3t + 2 = (t - 1)(t - 2) = 0$, 故齐次解为

$$y_H = Ae^{x} + Be^{2x},$$

这里 A 和 B 为未知常数. 由于微分方程的右边是 $7e^{2x}$, 由表可知特解的形式为 $y_P = Ce^{2x}$. 唉! 让人伤心的是, 这个选择会彻底失败. 事实上, 当令 $A = 0$ 且 $B = C$ 时, y_P 包含在 y_H 中. 这意味着若将 $y_P = Ce^{2x}$ 代入微分方程, 左边将得到 0,(试试!) 故该解无效. 然而, 如表的最后一行所示, 我们需要引入 x 的幂来使该解起有效. 因此, 我们将采用 $y_P = Cxe^{2x}$. 现在来看会发生什么. 首先注意 $y_P' = 2Cxe^{2x} + Ce^{2x}$ 且 $y_P'' = 4Cxe^{2x} + 4Ce^{2x}$, 故将其代入前面的微分方程时, 可得

$$(4Cxe^{2x} + 4Ce^{2x}) - 3(2Cxe^{2x} + Ce^{2x}) + 2Cxe^{2x} = 7e^{2x}.$$

关于 xe^{2x} 的项完全消掉了, 留下 $Ce^{2x} = 7e^{2x}$. 因此 $C = 7$, 这意味着 $y_P = 7xe^{2x}$. 最后, 全解为 $y = y_H + y_P = Ae^{x} + Be^{2x} + 7xe^{2x}$.

看另一个例子. 要想解

$$y'' + 6y' + 9y = e^{-3x},$$

需比原来做更进一步的计算. 齐次方程 $y'' + 6y' + 9y = 0$ 有特征二次式 $t^2 + 6t + 9 = (t+3)^2$, 因此齐次解为 $y_H = Ae^{-3x} + Bxe^{-3x}$. 由于微分方程的右边是 e^{-3x}, 因而我们取 $y_P = Ce^{-3x}$. 这个解无效, 因为它包含在 y_H 中 (当 $A = C$ 且 $B = 0$ 时). 甚至 $y_P = Cxe^{-3x}$ 也无效, 因为它也包含在 y_H 中 (当 $A = 0$ 且 $B = C$ 时). 因此我们需要进一步乘以 x^2 并令 $y_P = Cx^2e^{-3x}$. 现在可求两次导得 $y_P' = 2Cxe^{-3x} - 3Cx^2e^{-3x}$ 和 $y_P'' = 2Ce^{-3x} - 12Cxe^{-3x} + 9Cx^2e^{-3x}$. (对其验证!) 将这些量代入原方程并验证化简后为 $2Ce^{-3e} = e^{-3x}$ 的任务留给你完成. 这意味着 $C = \frac{1}{2}$, 故微分方程的解为 $y = y_H + y_P = Ae^{-3x} + Bxe^{-3x} + \frac{1}{2}x^2e^{-3x}$, 其中 A 和 B 为常数.

30.4.8 IVP

我们来看如何处理涉及常系数线性微分方程的 IVP. 跟通常一样, 要解 IVP, 首先解微分方程, 然后运用初始条件求剩下的未知常数.

我们将 30.4.6 节的两个例子改成 IVP, 然后求解. 对第一个例子, 假设给定 $y' + 2y = 4x + \frac{1}{3}\sin(5x)$, $y(0) = -1$. 现在暂时忽略条件 $y(0) = -1$, 我们已经知道通解是

$$y = Ae^{-2x} + 2x - 1 - \frac{5}{87}\cos(5x) + \frac{2}{87}\sin(5x).$$

又因为 $y(0) = -1$, 意味着当 $x = 0$ 时, $y = -1$. 将其代入, 可得

$$-1 = Ae^{0} + 2(0) - 1 - \frac{5}{87}\cos(0) + \frac{2}{87}\sin(0) = A - 1 - \frac{5}{87}.$$

化简为 $A = 5/87$, 故 IVP 的解为

$$y = \frac{5}{87}e^{-2x} + 2x - 1 - \frac{5}{87}\cos(5x) + \frac{2}{87}\sin(5x).$$

没有未知常数.



为了修改第二个例子, 我们假设 $y''-5y'+6y=2x^2\mathrm{e}^x$, $y(0)=y'(0)=0$. 如我们在 30.4.6 节所见, 通解 (忽略初始条件 $y(0)=0$ 和 $y'(0)=0$) 为

$$y=A\mathrm{e}^{2x}+B\mathrm{e}^{3x}+\left(x^2+3x+\frac{7}{2}\right)\mathrm{e}^x.$$

我们需将该解进行一次求导运算来运用初始条件 $y'(0)$ 的值, 验证

$$y'=2A\mathrm{e}^{2x}+3B\mathrm{e}^{3x}+\left(x^2+5x+\frac{13}{2}\right)\mathrm{e}^x.$$

因此, 当 $x=0$ 时, 我们知道 y 和 y' 都等于 0, 代入关于 y 的方程可得

$$0=A\mathrm{e}^0+B\mathrm{e}^0+\left(0^2+3(0)+\frac{7}{2}\right)\mathrm{e}^0=A+B+\frac{7}{2};$$

而代入关于 y' 的方程可得

$$0=2A\mathrm{e}^0+3B\mathrm{e}^0+\left(0^2+5(0)+\frac{13}{2}\right)\mathrm{e}^0=2A+3B+\frac{13}{2}.$$

同时解这些方程, 得到 $A=-4$ 和 $B=\dfrac{1}{2}$. 这意味着 IVP 的解为

$$y=-4\mathrm{e}^{2x}+\frac{1}{2}\mathrm{e}^{3x}+\left(x^2+3x+\frac{7}{2}\right)\mathrm{e}^x.$$

注意, 这两个例子都没有未知常数：初始条件使得我们能够求得唯一的解. 没有初始条件, 则总会有一个或两个未知常数.



我们来看最后一个 IVP 例子. 假设

$$y''+6y'+13y=26x^3-3x^2-24x,\quad y(0)=1,\quad y'(0)=2.$$

齐次方程是 $y''+6y'+13y=0$, 特征二次方程为 $t^2+6t+13=0$. 运用二次公式, 后面方程的解为 $t=(-6\pm\sqrt{36-4\cdot 13})/2=-3\pm 2\mathrm{i}$. 这意味着 $y_H=\mathrm{e}^{-3x}(A\cos(2x)+B\sin(2x))$. 现在来看特解：由于原方程的右边 (含 x 部分) 是三次的, 所以应写为 $y_P=ax^3+bx^2+cx+d$. 现在需要将 y_P 带入微分方程来求出从 a 到 d 的所有常数. 注意 $y_P'=3ax^2+2bx+c$ 和 $y_P''=6ax+2b$. 代入, 可得



$$(6ax+2b)+6(3ax^2+2bx+c)+13(ax^3+bx^2+cx+d)=26x^3-3x^2-24x.$$

令 x^3、x^2、x 和 1 的系数相等 (如我们在部分分式中所做一样), 分别可得 $13a=26$, $18a+13b=-3$, $6a+12b+13c=-24$, $2b+6c+13d=0$. 我把解这些方程的任务留给你完成, 可知 $a=2$, $b=-3$, $c=0$, $d=6/13$. 故 $y_P=2x^3-3x^2+6/13$, 因此

$$y=y_H+y_P=\mathrm{e}^{-3x}(A\cos(2x)+B\sin(2x))+2x^3-3x^2+\frac{6}{13},$$

对某些常数 A 和 B 成立. 为了求这些常数, 要使用初始条件. 由于 $y(0)=1$, 我们知道当 $x=0$ 时 $y=1$, 代入得到

$$1 = \mathrm{e}^{-3(0)}(A\cos(0) + B\sin(0)) + 2(0)^3 - 3(0)^2 + \frac{6}{13} = A + \frac{6}{13},$$

所以 $A = 7/13$. 同时, 对 y 的表达式求导得

$$y' = \mathrm{e}^{-3x}(-2A\sin(2x)) + 2B\cos(2x) - 3\mathrm{e}^{-3x}(A\cos(2x) + B\sin(2x)) + 6x^2 - 6x.$$

由于 $y'(0) = 2$, 可知当 $x = 0$ 时 $y' = 2$, 代入上面 y' 的表达式得到

$$\begin{aligned} 2 &= \mathrm{e}^0(-2A\sin(0) + 2B\cos(0)) - 3\mathrm{e}^0(A\cos(0) + B\sin(0)) + 6(0)^2 - 6(0) \\ &= 2B - 3A. \end{aligned}$$

因为 $A = 7/13$, 我们可以解最后这个方程来求出 $B = 47/26$. 将这些值代入, 求得最终解:

$$y = \mathrm{e}^{-3x}\left(\frac{7}{13}\cos(2x) + \frac{47}{26}\sin(2x)\right) + 2x^3 - 3x^2 + \frac{6}{13}.$$

注意, 该解中没有常数: 初始条件 (即 $y(0)$ 和 $y'(0)$ 的值) 确定了显式解.

30.5 微分方程建模

现实世界的很多量都可以用微分方程模拟 (即理论近似). 例如热流、波高、通货膨胀、电路的电流以及种群增长, 这些还只是一小部分. 下面是一个现实情形中涉及种群增长的简单例子.



某细菌培养以这样的方式呈指数增长: 它每小时的瞬时增长率等于培养皿中细菌数量的两倍. 假设某抗生素以每小时 8 盎司的恒定速率连续注入培养皿. 每盎司抗生素每小时杀死 25 000 细菌. 为保证培养皿细菌数量不为 0, 其初始数量至少需为多少?

这里的问题是: 随着细菌的繁殖, 细菌的数量在不断增长; 但随着抗生素不断注入培养皿, 抗生素的量也在增长. 哪个会赢, 细菌还是抗生素? 要解决这个问题, 我们需写出一个模拟该情形的微分方程. 事实上, 我们需要将文字问题转换成一个微分方程. 若没有抗生素, 则有标准种群增长微分方程为 ($k = 2$)

$$\frac{\mathrm{d}P}{\mathrm{d}t} = 2P,$$

其中 P 是在 t 小时时的种群数量. (我们在 9.6.1 节讨论过这类问题.) 现在我们需要把抗生素考虑进来, 修改方程. 在 t 小时, 我们知道有 $8t$ 盎司的抗生素, 所以细菌死亡率为 $8t \times 25\,000 = 200\,000t$. 因此正确的微分方程是

$$\frac{\mathrm{d}P}{\mathrm{d}t} = 2P - 200\,000t.$$

它可整理成标准形式

$$\frac{\mathrm{d}P}{\mathrm{d}t} - 2P = -200\,000t.$$

这个一阶线性方程的积分因子 (见 30.3 节) 是 $e^{\int -2\mathrm{d}t}$, 可化简为 e^{-2t}. 方程乘以积分因子, 得

$$e^{-2t}\frac{\mathrm{d}p}{\mathrm{d}t} - 2e^{-2t}P = -200\,000e^{-2t}t.$$

跟通常一样, 左边化简成 P 与积分因子之积的导数：

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{-2t}P) = -200\,000e^{-2t}t,$$

或直接写出

$$e^{-2t}P = -200\,000\int e^{-2t}t\mathrm{d}t.$$



右边需分部积分 (见 18.2 节), 证明

$$e^{-2t}P = 100\,000te^{-2t} + 50\,000e^{-2t} + 200\,000C,$$

这留给你来完成. 现在我们可以将 $200\,000C$ 替换为等价的任意常数 C. 同乘 e^{2t} 可得

$$P = 100\,000t + 50\,000 + Ce^{2t}.$$

这是时间为 t 的种群数量方程. 若初始数量为 P_0, 则可令方程中的 $t = 0$, 得到

$$P_0 = 100\,000(0) + 50\,000 + Ce^{2(0)} = 50\,000 + C.$$

这意味着 $C = P_0 - 50\,000$, 因此我们可将其代入方程, 得到

$$P = 100\,000t + 50\,000 + (P_0 - 50\,000)e^{2t}.$$

太好了! 我们掌握了关于这个情形的很多信息. 我们还需回答给定的问题. 当 P_0 为何值时会导致细菌数量最终为 0? 似乎 50 000 是一个临界值. 事实上, 若 $P_0 = 50\,000$, 上述方程就是 $P = 100\,000t + 50\,000$. 这种情况下, 细菌的初始数量为 50 000, 并以恒定的速率每小时 100 000 增长, 因此细菌永远不会灭绝. 若 $P_0 > 50\,000$, 则要加上 e^{2t} 的正数倍, 细菌数量增长得更快. 若 $P_0 < 50\,000$ 呢? 此时 $P_0 - 50\,000$ 是负的, 故我们有

$$P = 100\,000t + 50\,000 + (\text{负常数})e^{2t}.$$

因为最终指数起决定作用, 显然若 t 足够大, P 最终趋于 0. 例如, 即使初始细菌数量为 49 999, 我们有

$$P = 100\,000t + 50\,000 - e^{2t}.$$

图 30-1 是该情形下的 P-t 图.

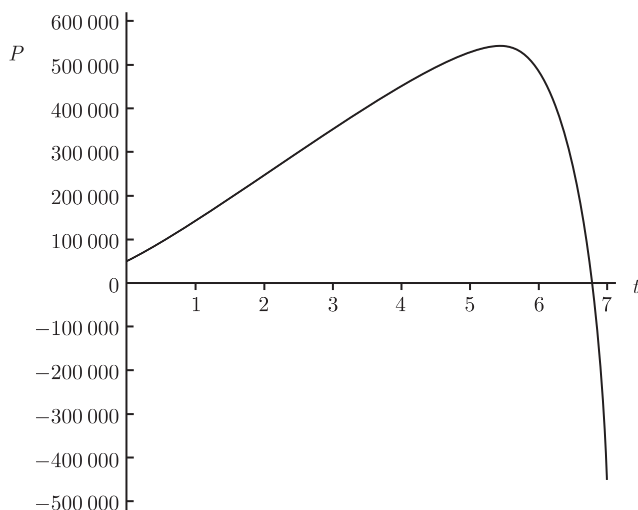


图 30-1

可以看到, 在前 5 个小时细菌数量近乎线性增长, 然后有一个快速的逆转, 最后在 6.5 和 7 小时之间的某处等于 0. (当然, 一旦等于 0, 讨论就结束了 —— 细菌数量永远不会小于 0, 因为细菌数量不能为负! 所以图中并未准确地反映 $P < 0$ 的情形.) 我们由此得出的一般结论是：若初始细菌数量小于 50 000, 则细菌将灭绝; 而若初始数量为 50 000 或更多, 培养皿内的细菌会幸存下来. 事实上, 它将持续增长.

附录 A　极限及其证明

贯穿本书, 我们使用了大量的极限. 极限很重要, 它也是导数定义和积分定义的核心部分. 正因如此, 我们是该以适当方式来定义它们了. 一旦我们知道它们是如何起作用的, 就可以证明许多原以为理所当然的事实. 以下就是本附录的内容：

- 极限的正式定义 (包括左极限与右极限、无穷极限、在 $\pm\infty$ 处的极限及数列的极限);
- 联合极限及三明治定理的证明;
- 连续和极限的关系, 包括介值定理的证明;
- 微分和极限, 包括乘积法则、商法则及链式求导法则的证明;
- 有关分段函数结果的证明及其导数;
- e 的存在性证明;
- 极值定理、罗尔定理、中值定理 (对于导数)、线性化中的误差公式及洛必达法则的证明;
- 泰勒近似定理的证明.

A.1　极限的正式定义

我们从函数 f 和实数 a 开始. 在 3.1 节中, 我们引入了记号

$$\lim_{x\to a} f(x) = L,$$

它贯穿整本书. 直观上, 上述方程意味着, 当 x 接近于 a 时, $f(x)$ 的值就会极度接近 L. 但有多近呢？想多近就有多近. 要了解这意味着什么, 让我们来做个小游戏.

A.1.1　小游戏

以下就是游戏规则. 你需要在 y 轴上选择一个以 L 为中点的区间并在其中移动, 画平行于 x 轴且通过区间端点的线. 如图 A-1 所示.

注意, 我用 $L-\varepsilon$ 和 $L+\varepsilon$ 标记了该区间的端点, 故两个端点到 L 的距离都是 ε.

不管怎样, 关键是不允许该函数的任意部分落在那两条水平线之外. 那么, 我的移动就是通过限制定义域来舍弃该函数的某些部分. 我只需要确保新的定义域是一个以 a 为中心的区间, 且该函数的每一点都位于你的两条线之间, 可能 $x = a$ 时除外. 图 A-2 是我的移动, 这基于你刚才的移动. 我可以舍弃更多的函数部分, 这当然没有问题, 只要剩余部分在那两条线之间就行了.



图　A-1

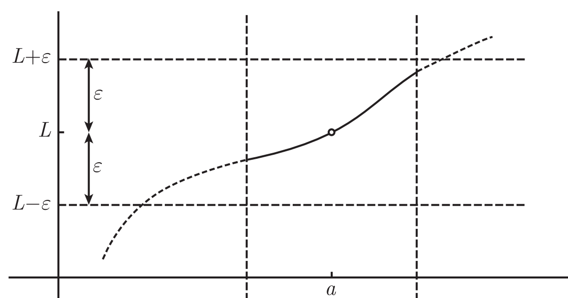


图　A-2

现在, 该你移动了. 你已然意识到, 当你的那两条线彼此接近时, 我的任务就更艰难了. 因此, 这一次你选取了一个更小的 ε 值. 图 A-3 是你第二次移动之后的情况.

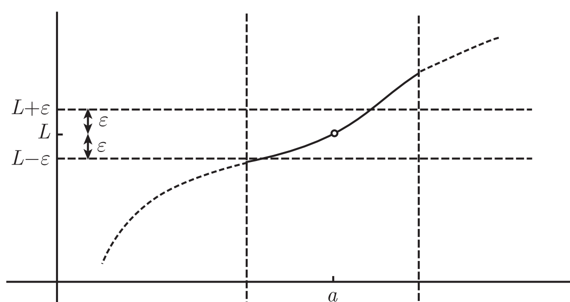


图　A-3

曲线上有一部分又落在两条水平线之外了, 可我还没移动呢. 我要舍弃更多的远离 $x = a$ 的函数部分, 如图 A-4 所示. 因此, 我又能够对抗你的移动了.

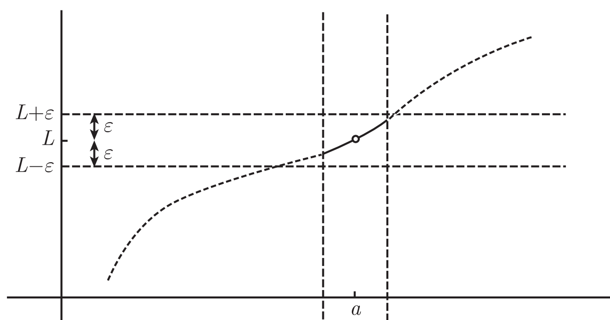


图　A-4

游戏何时结束呢？希望答案是游戏绝不停止! 不管你让那两条线多么接近, 只要我总是可以移动; 这实际上就是说 $\lim\limits_{x\to a} f(x)=L$ 是成立的. 我们不断缩小区间：你让那两条线不断接近, 我的回应是只关注函数足够接近 $x=a$ 的部分. 另一方面, 如果某次我的移动被卡住了, 那么 $\lim\limits_{x\to a} f(x)=L$ 就不再成立了. 极限或许是其他的值, 或者不存在, 但它一定不是 L.

A.1.2　真正的定义

我们需要将这个游戏转变为更多的符号. 首先注意你选择的区间是 $(L-\varepsilon,\ L+\varepsilon)$. 事实上, 你也可以将这个区间看作是满足 $|y-L|<\varepsilon$ 的点 y 的集合. 为什么呢？因为 $|y-L|$ 就是在数轴 (如 y 轴) 上 y 和 L 之间的距离. 因此, 你的区间是由所有距离 L 小于 ε 的点组成的. 正如你猜测的, 能够将 $|y-L|<\varepsilon$ 这样的不等式与其等价形式 $L-\varepsilon<y<L+\varepsilon$ 相互转换, 对于你来说是极其有帮助的.

现在轮到我移动了. 我需要保证该函数落在你的区间里. 这意味着, 在我舍弃大部分定义域之后, 所有保留下来的 $f(x)$ 的值都必须距离 L 小于 ε. 因此, 在我移动之后, 你将得出结论

$$|f(x)-L|<\varepsilon\ (x\text{ 充分接近于 }a\text{ 且 }x\neq a).$$

为了让我的移动更精确, 除了那个以 a 为中心的区间, 剩下的一切我都要舍弃. 我的区间看起来像是对某个其他数 δ 成立的 $(a-\delta,\ a+\delta)$, 因此, 我也可以把它看作是使得 $|x-a|<\delta$ 成立的 x 的集合. 事实上, 由于我不想让 x 等于 a, 所以可以写成 $0<|x-a|<\delta$.

总的来说, 你的移动是由选取 $\varepsilon>0$ 构成的.(它最好是正的, 否则根本不存在移动区间!) 而我的移动是选取一个数 $\delta>0$, 使得

$$|f(x)-L|<\varepsilon\ (0<|x-a|<\delta).$$

这意味着只要 x 距离 $a(x\neq a)$ 不超过 δ, $f(x)$ 的值距离 L 就不会超过 ε. 这就确定了基本思想：当 x 接近 a 时, $f(x)$ 接近 L. 现在, 剩下的就是允许你来选择 ε, 你

想要多小就多小, 而我仍然需要相应地选取 δ. 以下就是我们要找的正式定义：

> “$\lim_{x\to a} f(x) = L$” 表示, 对于任选的 $\varepsilon > 0$, 可以选取 $\delta > 0$, 使得：
> 对于所有满足 $0 < |x-a| < \delta$ 的 x, 有 $|f(x) - L| < \varepsilon$.

重要的是, 在你移动之后我才能开始移动! δ 的选取依赖于 ε 的选取. 通常我不能选择一个普遍的 δ 来保证对每一个 $\varepsilon > 0$ 结论都成立. 我必然受限于你的选择.

A.1.3 应用定义的例子



作为一个简单的例子 (不使用连续性), 我们来证明

$$\lim_{x\to 3} x^2 = 9.$$

我们很容易写出 $3^2 = 9$ 并宣布大功告成, 但这行不通, 因为极限只依赖于当 x 接近而不是等于 3 时的行为. 因此, 我们必须玩一下那个小游戏. 你选择 $\varepsilon > 0$, 这就产生了一扇约束我的小窗 $(9-\varepsilon,\ 9+\varepsilon)$. 现在, 我开始选取 δ. 假设 ε 是 8 (从上下文来看它过大了), 那么你的小窗是 $(1,\ 17)$. 现在, 通过选择 $\delta = 1$, 我可以轻易地留在那里, 此时我的小窗是 $(2,\ 4)$. (请记住, 我的小窗中心位于 3, 而你的小窗中心位于 9) 事实上, 如果你对介于 2 和 4 之间的任意数取平方, 就会得到一个介于 4 和 16 之间的数, 因此我的移动不成问题. 如果 ε 比 8 还大, 这会加宽你的区间, 若我坚持 $\delta = 1$, 那我的移动不成问题.

现在, 如果你选择的容忍度 ε 小于 8, 我就必须改变策略. 在这种情况下, 我的选择将是 $\delta = \varepsilon/8$. 即, 不管你如何选择, 我的小窗都将是你的小窗的 1/8. 要想知道这是怎么回事, 就得聪明点. 基本上, 我们必须选取在我区间内的任意一个数, 再平方, 并证明它位于你的区间内. 我的区间是 $(3-\varepsilon/8,\ 3+\varepsilon/8)$, 而你的区间是 $(9-\varepsilon,\ 9+\varepsilon)$.

让我们在我的区间中选取 x. 它能有多大呢?它必须小于 $3+\varepsilon/8$. 即 $x < 3+\varepsilon/8$, 也可以将它写为 $x-3 < \varepsilon/8$. 顺便说的是, 由于你的 ε 小于 8, 故我的 x 小于 4. 因此, 使用这两个不等式 $x-3 < \varepsilon/8$ 及 $x < 4$, 我们得到

$$(x-3)(x+3) < \left(\frac{\varepsilon}{8}\right)(4+3) = \frac{7\varepsilon}{8}.$$

由于 $(x-3)(x+3)$ 正好是 x^2-9, 所以我们可以在方程两边加上 9, 得到

$$x^2 < 9 + \frac{7\varepsilon}{8}.$$

故容忍上限 (两条线中上边的一条) 没有问题. 我们需要 $x^2 < 9+\varepsilon$, 刚刚证明过了. 那么容忍下限如何呢？那好, 已知 x 位于我的区间 $(3-\varepsilon/8, 3+\varepsilon/8)$ 中, 那么它能有多小呢？它必然大于 $3-\varepsilon/8$, 因此我们有 $x > 3-\varepsilon/8$; 这意味着, $x-3 > -\varepsilon/8$. 因为你的 ε 小于 8, 也有 $x-3 > -8/8 = -1$, 这就是说 $x > 2$. 现在应用不等式

$x-3>-\varepsilon/8$ 和 $x>2$, 我们得到

$$(x-3)(x+3)>\left(-\frac{\varepsilon}{8}\right)(2+3)=-\frac{5\varepsilon}{8}.$$

再次使用 $(x-3)(x+3)=x^2-9$, 我们在方程两边加上 9, 得到

$$x^2>9-\frac{5\varepsilon}{8}.$$

这样就有了容忍下限! 我们已经证明了, 如果 x 位于区间 $(3-\varepsilon/8, 3+\varepsilon/8)$ 内, 那么 x^2 就在区间 $(9-5\varepsilon/8,\ 9+7\varepsilon/8)$ 内. 由于 5/8 和 7/8 都小于 1, 所以我们可以确信 x^2 位于区间 $(9-\varepsilon,\ 9+\varepsilon)$ 内; 毕竟, 这个区间包含前面那个.

综合起来, 我们设 $f(x)=x^2$, 且要证明

$$\lim_{x\to 3}f(x)=9.$$

你选择 ε, 而我就相应地选取 $\delta=\varepsilon/8$, 除非你的 ε 比 8 大, 若如此, 我就选取 $\delta=1$. 我们已经证明了, 在这两种情况下, 如果 x 位于区间 $(3-\delta,\ 3+\delta)$ 内, 那么 $f(x)$ 就在区间 $(9-\varepsilon,\ 9+\varepsilon)$ 内. 换句话说, 只要 $|x-3|<\delta$, 那么 $|f(x)-9|<\varepsilon$. 如果指明 $0<|x-3|<\delta$, 那么 $|f(x)-9|<\varepsilon$ 的话, 我们也可以把 $x=3$ 排除在外. 这正是我们想要的 —— 证明了上述等式. 信不信由你, 如果想要利用定义来证明以上极限成立, 那么必须做大量的工作!

A.2　由原极限产生新极限

前面的例子令人十分烦恼. 仅仅是想证明当 $x\to 3$ 时 $x^2\to 9$, 我们就必须做大量的工作. 幸运的是, 事实表明, 一旦你知道一些极限, 就可以将它们放在一起讨论并得到一大堆新的极限. 例如, 你可以在合理的范围内对极限做加法、减法、乘法及除法, 也可以使用三明治定理. 下面就让我们来看看为什么这些都是成立的.

A.2.1　极限的和与差及证明

假设我们有两个函数 f 和 g, 并且知道, 当 $x\to a$ 时 $f(x)\to L$ 和 $g(x)\to M$. 那么, 当 $x\to a$ 时, $f(x)+g(x)$ 会怎样呢? 直观上, 它应该是趋于 $L+M$ 的. 让我们用定义来证明它. 我们知道

$$\lim_{x\to a}f(x)=L \quad 和 \quad \lim_{x\to a}g(x)=M.$$

这意味着, 如果你选取 $\varepsilon>0$, 我可以将 x 限制为充分接近 a 来保证 $|f(x)-L|<\varepsilon$. 如果 x 充分地接近 a, 我也可以保证 $|g(x)-M|<\varepsilon$. 对于 f 和 g 来说, 我需要的接近程度或许是不同的, 但这没有问题 —— 我可以做到充分接近, 以便两个不等式都成立.

如果 $f(x)+g(x)$ 接近 $L+M$, 则这两个量之间的差异应该很小. 因此我们需要查看量 $|(f(x)+g(x))-(L+M)|$. 我们将它写为 $|(f(x)-L)+(g(x)-M)|$. 然

后, 我们使用所谓的三角不等式 (就是说[①] 对于任意的数 a 和 b, 有 $|a+b| \leqslant |a|+|b|$) 得到

$$|(f(x)-L)+(g(x)-M)| \leqslant |f(x)-L|+|g(x)-M| < \varepsilon+\varepsilon=2\varepsilon,$$

这里假设 x 充分地接近 a. 这已经够好了, 不过你想要的容忍限度是 ε, 而不是 2ε! 因此, 我必须再次移动 (对不起了); 这一次, 我要将小窗变窄, 以便 $|f(x)-L|$ 和 $|g(x)-M|$ 都小于 $\varepsilon/2$, 而不是 ε. 这是没有问题的, 因为我可以应对你选取的任意一个正数. 不管怎样, 如果你再做一遍上述方程的话, 在右边将得到 ε 而不是 2ε, 这样, 我们就证明了可以找到一扇关于 a 的小窗使得

$$|(f(x)+g(x))-(L+M)| < \varepsilon$$

在此假设 x 在我的小窗里. (如果你想要将这扇小窗描述得更好, 也可以使用 δ, 但事实上它不会给我们提供任何附加信息.) 因此, 这就证明了

$$\text{如果} \lim_{x\to a} f(x)=L \text{且} \lim_{x\to a} g(x)=M, \text{那么} \lim_{x\to a}(f(x)+g(x))=L+M.$$

即, 和的极限等于极限的和. 它的另一种写法是

$$\lim_{x\to a}(f(x)+g(x))=\lim_{x\to a} f(x)+\lim_{x\to a} g(x).$$

但在这里, 你必须非常仔细检验, 以确保右边的这两个极限是存在且有限的. 如果其中一个极限是 $\pm\infty$ 或不存在, 就不能再运用上述公式了. 两个极限都必须是有限的, 也能确保可以相加. 如果它们不存在, 你或许很幸运, 但这没有保证.

$f(x)-g(x)$ 又如何呢? 它应该是趋于 $L-M$ 的. 确实如此:

$$\text{如果} \lim_{x\to a} f(x)=L \text{且} \lim_{x\to a} g(x)=M, \text{那么} \lim_{x\to a}(f(x)-g(x))=L-M.$$



该证明几乎和我们刚刚看到的那个差不多, 不过你需要一个略有不同的三角不等式: $|a-b| \leqslant |a|+|b|$. 事实上, 这就是应用于 a 和 $-b$ 的三角不等式, 即 $|a+(-b)| \leqslant |a|+|-b|$, 当然有 $|-b|$ 等于 $|b|$. 现在就由你来重新写出以上论证, 但要将 $f(x)$ 和 $g(x)$ 以及 L 和 M 之间的加号改为减号.

A.2.2 极限的乘积及证明

现在, 我们再来假设两个函数 f 和 g 满足

$$\lim_{x\to a} f(x)=L \quad \text{且} \quad \lim_{x\to a} g(x)=M.$$

我们想证明

$$\lim_{x\to a} f(x)g(x)=LM.$$

---

① 既然我们在证明各种命题, 那不妨也证明一下三角不等式. 首先我们要注意到, 对任意数 x 都有 $x \leqslant |x|$. 实际上, 若 x 为正数或等于 0, 则 $x=|x|$; 若 x 为负数, 则由 $|x|$ 为正数可得 $x<|x|$. 现在用 ab 替换 x 就得到 $ab \leqslant |ab|=|a|\cdot|b|$. 两边先同时乘以 2 再加上 a^2+b^2 可得 $a^2+b^2+2ab \leqslant a^2+b^2+2|a|\cdot|b|$. 左边就是 $(a+b)^2$. 因为对任意的 x 都有 $x^2=|x|^2$, 所以左边可写作 $|a+b|^2$. 同样地, 右边可写作 $|a|^2+|b|^2+2|a|\cdot|b|$, 或 $(|a|+|b|)^2$. 于是就有不等式 $|a+b|^2 \leqslant (|a|+|b|)^2$. 现在我们两边同时开方, 因为 $|a+b|$ 和 $|a|+|b|$ 都非负, 于是就得到了三角不等式.

即, 乘积的极限等于极限的乘积. 它的另一种写法是

$$\lim_{x\to a} f(x)g(x) = \lim_{x\to a} f(x) \times \lim_{x\to a} g(x).$$

我们同样要知道的是, 右边的这两个极限是存在的且为有限的. 为了求证, 我们需要证明 $f(x)g(x)$ 与 (希望中的) 极限 LM 的差是很小的. 我们来考虑差 $f(x)g(x) - LM$. 技巧是减去 $Lg(x)$ 再加上它! 即,

$$f(x)g(x) - LM = f(x)g(x) - Lg(x) + Lg(x) - LM.$$

我们会得到什么呢? 我们来取绝对值, 然后使用三角不等式:

$$\begin{aligned}|f(x)g(x) - LM| &= |(f(x) - L)g(x) + L(g(x) - M)| \\ &\leqslant |(f(x) - L)g(x)| + |L(g(x) - M)|.\end{aligned}$$

整理一下可写为

$$|f(x)g(x) - LM| \leqslant |f(x) - L| \cdot |g(x)| + |L| \cdot |g(x) - M|.$$

现在, 该玩游戏了. 你选取你的正数 ε, 然后我开始工作. 我将关注环绕 $x = a$ 的极小区间, 以便 $|f(x) - L| < \varepsilon$ 且 $|g(x) - M| < \varepsilon$. 事实上, 如果你选取 $\varepsilon \geqslant 1$(一个十分无力的移动, 因为你想要 ε 非常小!), 那么我甚至会继续坚持 $|g(x) - M| < 1$. 因此, 我们知道, 不管在哪种情况下都有 $|g(x) - M| < 1$; 这意味着, 在我的区间上 $M - 1 < g(x) < M + 1$. 特别地, 我们可以看到 $|g(x)| < |M| + 1$. 要点在于, 在我的区间上有一些理想的不等式:

$$|f(x) - L| < \varepsilon, \quad |g(x)| < |M| + 1, \quad |g(x) - M| < \varepsilon.$$

我们可以将它们代入上面的 $|f(x)g(x) - LM|$, 得到

$$\begin{aligned}|f(x)g(x) - LM| &\leqslant |f(x) - L| \cdot |g(x)| + |L| \cdot |g(x) - M| \\ &< \varepsilon \cdot (|M| + 1) + |L| \cdot \varepsilon = \varepsilon(|M| + |L| + 1),\end{aligned}$$

其中 x 充分地接近 a. 这几乎就是我们想要的了! 在右边我应该得到 ε, 但是我得到了一个额外的因子 $(|M| + |L| + 1)$. 这没有问题 —— 只要你允许我再移动一次, 这一次我将确保 $|f(x) - L|$ 不超过 $\varepsilon/(|M| + |L| + 1)$. $|g(x) - M|$ 同理. 然后, 当我重做以上步骤时, ε 将由 $\varepsilon/(|M| + |L| + 1)$ 代替, 并且在最后一步, 因子 $(|M| + |L| + 1)$ 会被消除, 而我们正好得到 ε! 这样, 我们就证明了该结论.

顺便一提的是, 要注意上述情况的一个特例. 如果 c 是常数, 那么



$$\lim_{x\to a} cf(x) = c \lim_{x\to a} f(x).$$

在上述主要公式中设 $g(x) = c$, 很容易看出这一点. 我将细节留给你来完成.

A.2.3 极限的商及证明

现在, 我们重做一下练习. 如果

$$\lim_{x\to a} f(x) = L \quad 和 \quad \lim_{x\to a} g(x) = M,$$

那么, 我们有

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{L}{M}.$$

因此, 商的极限等于极限的商. 为了让它有意义, 我们要保证 $M \neq 0$, 否则就要除以 0 了. 以上等式的另一种写法是

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x \to a} f(x)}{\lim\limits_{x \to a} g(x)},$$

只要这两个极限都存在且为有限的, 同时 g 的极限非零.

以下是求证过程. 我们想要 $f(x)/g(x)$ 接近 L/M, 因此, 要考虑它们的差. 然后, 我们需要通分:

$$\frac{f(x)}{g(x)} - \frac{L}{M} = \frac{Mf(x) - Lg(x)}{Mg(x)}.$$

现在, 我们使用一个类似于极限的乘积中的技巧: 在分子上减去并加上 LM, 然后做因式分解, 会得到

$$\begin{aligned}\frac{f(x)}{g(x)} - \frac{L}{M} &= \frac{Mf(x) - LM + LM - Lg(x)}{Mg(x)} \\ &= \frac{M(f(x) - L)}{Mg(x)} + \frac{L(M - g(x))}{Mg(x)} \\ &= \frac{f(x) - L}{g(x)} - \frac{L(g(x) - M)}{Mg(x)}.\end{aligned}$$

如果我们取绝对值, 然后使用形如 $|a - b| \leqslant |a| + |b|$ 的三角不等式, 将得到

$$\left|\frac{f(x)}{g(x)} - \frac{L}{M}\right| = \left|\frac{f(x) - L}{g(x)} - \frac{L(g(x) - M)}{Mg(x)}\right| \leqslant \left|\frac{f(x) - L}{g(x)}\right| + \left|\frac{L(g(x) - M)}{Mg(x)}\right|.$$

因此, 你通过选取 $\varepsilon > 0$ 来移动, 然后我会将 $x = a$ 附近的那扇小窗变窄, 使得在这扇小窗中, $|f(x) - L| < \varepsilon$ 且 $|g(x) - M| < \varepsilon$. 现在, 我需要变得更加聪慧. 你看, 我知道 $M - \varepsilon < g(x) < M + \varepsilon$, 这表示 $|g(x)| > |M| - \varepsilon$. 如果右边的量 $|M| - \varepsilon$ 为正, 那么一切不成问题; 但是如果它是负的, 将不会告诉我们任何信息, 因为我们已经知道 $|g(x)|$ 不可能是负的. 因此, 如果 ε 足够小, 那么我就不担心了; 但是如果它大一点的话, 我就需要将我的那扇小窗变窄, 使得 $|g(x)| > |M|/2$. 总之, 在这个区间上, 我们有三个不等式成立:

$$|f(x) - L| < \varepsilon, \quad |g(x)| > \frac{|M|}{2}, \quad |g(x) - M| < \varepsilon.$$

中间的那个不等式颠倒过来为

$$\frac{1}{|g(x)|} < \frac{2}{|M|}.$$

综述, 我们有

$$\left|\frac{f(x)}{g(x)} - \frac{L}{M}\right| \leqslant \frac{|f(x) - L|}{|g(x)|} + \frac{|L| \cdot |g(x) - M|}{|M||g(x)|} < \varepsilon \cdot \frac{2}{|M|} + \varepsilon \cdot \frac{|L|}{|M|} \cdot \frac{2}{|M|}.$$

这还不是我们想要的——这里有一个额外的因子 $\left(2/|M| + 2|L|/|M|^2\right)$. 但我们知道如何处理它——只需要再移动一次, 这一次不针对 ε, 而是 ε 除以这个额外因子.

A.2.4 三明治定理及证明

在 3.6 节中, 我们见过三明治定理. 现在该证明它了. 我们以函数 f、g 和 h 开始, 它们满足对于所有充分接近 a 的 x, 有 $g(x) \leqslant f(x) \leqslant h(x)$. 我们也知道

$$\lim_{x\to a} g(x) = L \quad 和 \quad \lim_{x\to a} h(x) = L.$$

直观上, f 被越来越紧地夹在 g 和 h 之间, 以至于当 $x \to a$ 时, 我们应该会有 $f(x) \to L$. 即, 我们需要证明

$$\lim_{x\to a} f(x) = L.$$

好吧, 你开始选取你的正数 ε, 然后我关注一个中心位于 a 的小区间, 使此区间 $|g(x) - L| < \varepsilon$ 且 $|h(x) - L| < \varepsilon$. 我还需要不等式 $g(x) \leqslant f(x) \leqslant h(x)$ 在此区间成立; 由于不等式或许只有当 x 非常接近 a 时成立, 因而我可能必须缩减我的原始区间.

不管怎样, 我们知道当 x 充分接近 a 时, $|h(x) - L| < \varepsilon$; 该不等式可以重写为

$$L - \varepsilon < h(x) < L + \varepsilon.$$

事实上, 我们只需要右边的不等式 $h(x) < L + \varepsilon$. 你看, 在我的小区间里可知 $f(x) \leqslant h(x)$, 因此有

$$f(x) \leqslant h(x) < L + \varepsilon.$$

类似地, 我们知道

$$L - \varepsilon < g(x) < L + \varepsilon,$$

其中 x 充分接近 a. 这一次, 我们舍弃右边的不等式而使用 $g(x) \leqslant f(x)$ 会得到

$$L - \varepsilon < g(x) \leqslant f(x).$$

综述, 我们证明了, 当 x 接近 a 时,

$$L - \varepsilon < f(x) < L + \varepsilon,$$

或简单的形式 $|f(x) - L| < \varepsilon$. 这样, 我们就证明了三明治定理!

A.3 极限的其他情形

现在我们来快速看一些其他类型极限的定义: 无穷极限、左极限与右极限及在 $\pm\infty$ 处的极限.

A.3.1 无穷极限

使用我们的游戏来定义极限

$$\lim_{x\to a} f(x) = \infty$$

是不适用的. 当你尝试画出那两条接近极限的线时, 就会被完全卡住, 因为该极限应该是 ∞ 而不是某个有限的值 L. 因此, 我们必须对规则做些修正. 我的移动不会有太大改变, 但是你的移动会变化很大. 这一次你不是要选取一个很小的数 ε 再画

出两条水平线 (高度为 $L-\varepsilon$ 和 $L+\varepsilon$), 而是要选取一个很大的数 M 并且只画一条线, 其高度为 M. 我仍然会通过舍弃该函数的大部分来移动, 不过保留一个围绕 $x=a$ 的很小部分. 尽管如此, 这一次我必须确保所剩部分总是在你那条线的上方. 例如, 图 A-5 显示了你的一次移动及我接下来的反应.

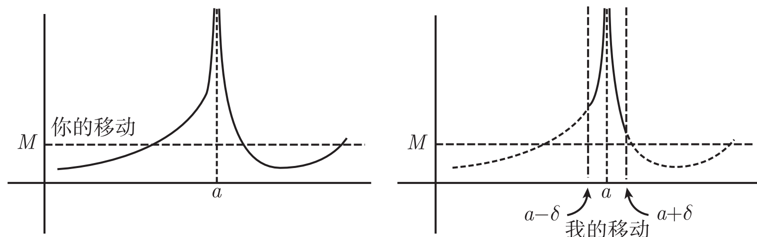


图 A-5

如果你做另一次移动, 这次 M 的值更大了, 会发生什么呢? (参见图 A-6.)

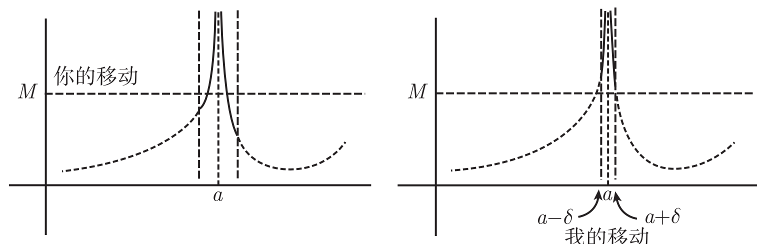


图 A-6

因此其基本思想就是, 这次你将那条线提升得越来越高了; 如果我总是可以对你的移动做出反应, 那么该极限的确是 ∞. 用符号表示就是, 我需要能够保证不管 M 有多大, 只要 x 充分接近 a 就有 $f(x)>M$. 定义如下:

> $\lim\limits_{x\to a} f(x)=\infty$ 表示对于你选取的任意的 $M>0$, 我都可以选取$\delta>0$, 使得:
> 对于所有满足 $0<|x-a|<\delta$ 的 x, $f(x)>M$.

这和极限是某个有限数 L 时的情况差不多, 只是不等式 $|f(x)-L|<\varepsilon$ 由 $f(x)>M$ 所代替.

例如, 假设我们想要证明

$$\lim_{x\to 0}\frac{1}{x^2}=\infty.$$

以你选取数 M 开始; 然后, 我必须确保当 x 充分接近 a 时 $f(x)>M$. 那好, 假设我舍弃满足 $|x|<1/\sqrt{M}$ 的 x 之外的一切. 对于这样的一个 x, 我们有 $x^2<1/M$, 故 $1/x^2>M$(注意我们假设了 $x\neq 0$). 这意味着在我的区间里 $f(x)>M$, 也就是



说我的移动是有效的. 对于你选取的任意 M, 我都可以做出一个有效的移动, 这样, 我就证明了该极限的确是 ∞.

那么 $-\infty$ 的情况会怎样呢？一切正好反转过来. 你仍然选取一个很大的正数 M, 但这一次, 我需要在移动的同时确保该函数总是在高度为 $-M$ 的水平线的下方. 故定义如下：

$\lim\limits_{x\to a} f(x) = -\infty$ 表示对于你选取的任意的$M > 0$,我都可以选取$\delta > 0$,使得：
对于所有满足 $0 < |x-a| < \delta$ 的 $x, f(x) < -M$.

A.3.2 左极限与右极限

为了定义右极限, 我们来做相同的游戏, 只是这一次我们提前舍弃了 $x = a$ 左边的一切. 其效果就是：当我移动时只需要考虑 $(a,\ a+\delta)$, 而不是选取一个类似 $(a-\delta,\ a+\delta)$ 的区间. a 左边的一切都是无关紧要的.

类似地, 对于左极限, 只有 a 左边的 x 值是相关的. 这表示, 我的区间会形如 $(a-\delta,\ a)$; 我已经舍弃了 $x = a$ 右边的一切.

这一切表明, 你可以取以上方框定义中的任意一个, 并将不等式 $0 < |x-a| < \delta$ 改为 $0 < x-a < \delta$ 来得到右极限. 而为了得到左极限, 就要将不等式 $0 < |x-a| < \delta$ 改为 $0 < a-x < \delta$. 你不用详细写出全部六种形式 (就是极限值为 L、∞ 及 $-\infty$ 的左极限和右极限), 但若你能不看这几页, 自己尝试全部写出, 那的确是一个不错的练习机会.

A.3.3 在 ∞ 及 $-\infty$ 处的极限

最后的极限情形是在 ∞ 或 $-\infty$ 而不是某个有限值 a 取极限. 因此, 我们想要定义

$$\lim_{x\to\infty} f(x) = L.$$

当然需要对游戏稍作改动, 我们已经知道如何去做了. 事实上, 只需要改写 A.3.1 节中的方法. 以你选取很小的数 $\varepsilon > 0$ 开始, 建立你的容忍区间 $(L-\varepsilon,\ L+\varepsilon)$; 然后我的移动将是舍弃某条垂线 $x = N$ 左侧的函数部分, 以便这条线右侧的所有函数值都落在你的容忍区间内; 然后, 你选取一个更小的 ε, 如果我必须落在你的那个新的小区间内, 就要将那条垂线右移. 图 A-7 所示就是我们两个开始的几次可能的移动.

在你第一次移动后, 我的移动保证了垂线 $x = N$ 右侧的函数值都落在你的容忍区间内. 你的反应是使两条水平线逼近, 而我只需要将垂线右移, 直到我可以满足你的限制更严的新容忍区间. 同样, 如果我总是可以对你的移动做出反应, 那么以上极限成立.

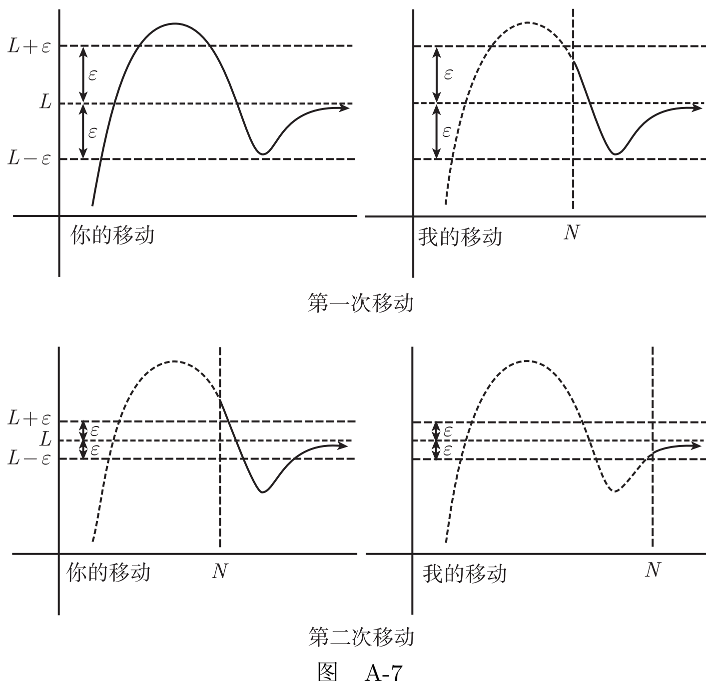


图 A-7

更正式地, 我的移动由选取 N 组成, 使得只要 $x>N$(x 位于垂线 $x=N$ 的右侧), 都有 $f(x)$ 位于区间 $(L-\varepsilon,\ L+\varepsilon)$. 使用绝对值, 我们可以将它写作：

> $\lim\limits_{x\to\infty} f(x)=L$ 表示对于你选取的任意 $\varepsilon>0$, 我都可以选取N,使得：
> 对于所有满足 $x>N$ 的 x, $|f(x)-L|<\varepsilon$.

需要注意的是, $x\to\infty$ 时的极限必定是一个左极限 —— 在 ∞ 的右侧什么也没有! 不管怎样, 我们仍然要看一些情形. 首先, $\lim\limits_{x\to\infty} f(x)=\infty$ 是什么意思? 你只需要改写之前的定义. 特别地, 你可以取上述定义并将你的移动变为选取 $M>0$, 现在用 $f(x)>M$ 代替 $|f(x)-L|<\varepsilon$. 反之, 如果你想要证明 $\lim\limits_{x\to\infty} f(x)=-\infty$, 就要将不等式改为 $f(x)<-M$. 这相当简单.

定义下列极限很简单：

$$\lim_{x\to-\infty} f(x)=L,\quad \lim_{x\to-\infty} f(x)=\infty,\quad \lim_{x\to-\infty} f(x)=-\infty$$

和 $x\to\infty$ 时的情形不同, 我的垂线变为 $x=-N$, 且现在的函数值都必须落在该线左侧你的容忍区间, 而不是右侧即你只需在所有的定义中将不等式 $x>N$ 改为 $x<-N$.

事实上, 我们可以使用相同的思想来定义无穷数列的极限. 在 22.1 节, 我们给出了一个非正式的定义, 现在可以做得更好. 我们由一个无穷数列 $a_1,a_2,a_3,\cdots$ 开

始, 那么

$$\lim_{n\to\infty} a_n = L \text{ 表示对于你选取的任意的 } \varepsilon > 0, \text{我都可以选取 } N, \text{使得：}\\ \text{对于所有满足 } n > N \text{ 的 } n, |a_n - L| < \varepsilon.$$

如果比较该定义和

$$\lim_{x\to\infty} f(x) = L$$

的定义, 你会看到它们几乎是一样的. 唯一的区别就是, 连续变量 x 由整数值型变量 n 代替. 此时, L 由 ∞ (或 $-\infty$) 代替, 然后你选择 $M > 0$ 而不是 $\varepsilon > 0$, 且不等式 $|a_n - L| < \varepsilon$ 变为 $a_n > M$ (或相应地有 $a_n < -M$).

现在, 如果你真的想要挑战一下, 就请尝试写出每一个可能的极限类型的定义吧. (我们看过 18 个!) 再来一次, 看你是否可以证明类似 A.2 节的所有结论在其他情况下的结果.

A.3.4　两个涉及三角函数的例子

在 3.4 节中, 我们说极限

$$\lim_{x\to\infty} \sin(x)$$

不存在 (DNE). 凭直觉, $\sin(x)$ 一直在 -1 和 1 之间振荡, 因此它不会趋于任意一个数. 我们现在用 A.3.3 节的定义来证明这个直觉是对的. 假设该极限存在且极限值为 L. 你选取数 $\varepsilon > 0$, 然后我需要选取一个很大的数 N, 只要 $x > N$, 就有 $|\sin(x) - L| < \varepsilon$. 我们假设你选取 ε 为 $\frac{1}{2}$. 这意味着, 我需要保证, 只要 $x > N$ 就有 $|\sin(x) - L| < \frac{1}{2}$. 从另一种方式来看, 就是对于所有的 $x > N$, $\sin(x)$ 必须落在区间 $\left(L - \frac{1}{2}, L + \frac{1}{2}\right)$ 中. 不幸的是, 不管 L 和 N 是什么, 这都是不可能的! 要知道为什么, 我们首先选取大于 N 的 π 的倍数：不妨设这个数为 $n\pi$. 其中 n 是一个整数. 那么, $\sin(n\pi + \pi/2) = 1$, 而 $\sin(n\pi + 3\pi/2) = -1$. $\sin(x)$ 的这两个值的距离为 2, 由于区间 $\left(L - \frac{1}{2}, L + \frac{1}{2}\right)$ 的长度仅为 1, 故它们不可能都落在该区间中. 因此, 该极限不可能是一个有限的数 L.

图 A-8 是我们对极限 L 期望的三个可能的候选图像.

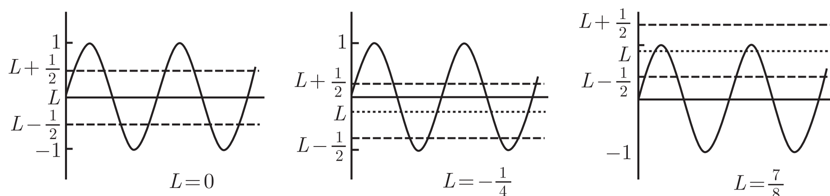


图　A-8

在每种情况下, 围绕 L 的区间的宽都是 $\frac{1}{2}$, 但是在这三种情况下, 即使我舍弃了函数的大部分, 还是不能将 $\sin(x)$ 填塞到该区间中. 由于 $\sin(x)$ 总超出区间, 而我总是在你的区间内移动, 因此我画不出一条垂线并说在该线的右侧. 不论对于哪个高度为 1 的水平线条, 结果都是一样的.

为了使讨论完整, 我们还应该确保该极限不可能是 ∞ 或 $-\infty$. 事实上, 如果该极限是 ∞ 的话, 那么你将选取 $M > 0$, 而我必须确保对于某个 N, 只要 $x > N$ 就有 $\sin(x) > M$. 然而, 要想阻挠我, 你只需选取 $M = 2$. 由于对于任意的 x 都不会有 $\sin(x) > 2$, 所以我就被钉住了. 可用相同的移动来处理 $-\infty$ 的情况 (做做看). 这样, 我们的确证明了以上极限不存在.



在 3.3 节中, 我们还提到极限

$$\lim_{x\to 0^+} \sin\left(\frac{1}{x}\right)$$

不存在. 为了证明这是真的, 你可以选取一个可能的极限 L 并进行如同上例的论证. 如果你的移动是选取 $\varepsilon = \frac{1}{2}$, 那么我需要试着选取 $\delta > 0$, 使得只要 $0 < x < \delta$ 就有 $|\sin(1/x) - L| < \frac{1}{2}$. (这里我们使用的是 A.3.2 节中的定义.) 现在你可以变聪明些并尝试找到两个会把上述情形搞乱的很小的 x 值. 事实上, 对于足够大的 n, 如果你尝试 $x = 1/(n\pi + \pi/2)$, 然后尝试 $x = 1/(n\pi + 3\pi/2)$, 那么两个取值都在 $0 < x < \delta$ 中, 但事实表明, $\sin(1/x)$ 的结果分别是 1 和 -1. 不管 L 如何, 它们两个不可能都落在容忍区间 $\left(L - \frac{1}{2}, L + \frac{1}{2}\right)$ 中, 这就说明 L 不是极限.



你应该尝试写出这些细节, 但有一个更简单的方法. 由于已知 $\lim\limits_{x\to\infty} \sin(x)$ 不存在, 所以可以只做简单的极限变量替换. 事实上, 如果令 $u = 1/x$, 那么 $x = 1/u$, 则我们立刻知道

$$\lim_{1/u\to\infty} \sin\left(\frac{1}{u}\right)$$

不存在. $1/u \to \infty$ 何时为真呢? 唯一的情况就是当 $u \to 0^+$. 一般来说, 证明这个切换并不难 (见 A.4.1 节), 因此, 我们看到

$$\lim_{u\to 0^+} \sin\left(\frac{1}{u}\right) \text{不存在.}$$

现在只需要将虚拟变量 u 改为 x, 不用费劲就会得到我们想要的结果了!

A.4 连续与极限

正如我们在 5.1.1 节中看到的, 说一个函数 f 在 $x = a$ 上连续就是指

$$\lim_{x\to a} f(x) = f(a).$$

即, 当 $x \to a$ 时有 $f(x) \to f(a)$. 因此, 函数 f 保持极限, 这就是连续的核心思想. 不管怎样, 现在我们可以使用极限的知识来证明, 当你对两个在 $x = a$ 上连续的函

数做加法、减法、乘法或除法时, 新的函数在那里也是连续的. (在除法的情况下, 分母在 $x=a$ 上不能为 0.) 事实上, 我们假设 f 和 g 在 $x=a$ 上连续, 那么

$$\lim_{x\to a} f(x)=f(a) \quad 和 \quad \lim_{x\to a} g(x)=g(a).$$

因此, 为了证明函数 $f+g$ 在 $x=a$ 上连续, 我们所要做的就是拆分极限. 在 A.2.1 节我们证明过

$$\lim_{x\to a}(f(x)+g(x))=\lim_{x\to a} f(x)+\lim_{x\to a} g(x)=f(a)+g(a).$$

就是这么简单. 现在, 你可以用 $-$、$\times$ 或 / 号来替换 $+$ 号, 从而得到对于减法、乘法和除法的类似结果.

A.4.1 连续函数的复合

我们来看一些稍复杂情况. 假设 f 和 g 都处处连续, 我们想要证明复合函数 $f\circ g$ 也处处连续. 我们需要集中考虑一个特殊的 x 值. 因此, 假设 g 在 $x=a$ 上连续. 那么我们需要 f 在哪里连续呢? 我们想要证明

$$\lim_{x\to a} f(g(x))=f(g(a)),$$

因此没有必要去担心 f 在 $x=a$ 上是否连续. 我们需要的是它在 $g(a)$ 上连续, 因为我们要在 $g(a)$ 的附近且在点 $g(a)$ 上评估 f.

下面就是我们面临的情况: g 在 $x=a$ 上连续, 且 f 在 $x=g(a)$ 上连续, 要证明 $f\circ g$ 在 $x=a$ 上连续. 为了求证, 我们需要在游戏中增加第三参与者. 事实上, 我将对抗这个新的参与者, 我们称之为 Smiddy, 而 Smiddy 将对抗你.

来看看如何玩游戏吧. 由于 f 在 $g(a)$ 上连续, 我们知道

$$\lim_{y\to g(a)} f(y)=f(g(a)).$$

注意, 我使用 y 作为代替 x 的虚拟变量, 但这没问题 —— 你可以将 y 变成你喜欢的任意字母, 它们表示的是同一个意义. 不管怎样, 我们设 $L=f(g(a))$. 然后, 你选取你的 $\varepsilon>0$, 建立你的容忍区间 $(L-\varepsilon,\ L+\varepsilon)$. 而你要挑战 Smiddy, 舍弃以 $y=g(a)$ 为中心的一个小区间外面的一切, 以便所剩的函数值都落在你的区间内. 即, Smiddy 应该选取 $\lambda>0$, 使得 $|y-g(a)|<\lambda$ 时都有 $|f(y)-L|<\varepsilon$. 因为以上极限是正确的, 所以 Smiddy 就可以这样做. 为什么要用 λ 代替 δ 呢? 因为 Smiddy 非常喜欢它.

现在, 轮到我来对抗 Smiddy 了. 这一次, 我们根据 g 在 $x=a$ 上连续的事实写出

$$\lim_{x\to a} g(x)=g(a).$$

关键是: Smiddy 使用的是数 λ, 而不是你已经使用的 ε! 因此, Smiddy 的容忍区间是 $(g(a)-\lambda,\ g(a)+\lambda)$. 现在, 我必须舍弃以 $x=a$ 为中心的一个小区间外的一切, 以便所剩的函数值落在 Smiddy 的区间内. 因为以上极限是正确的, 所以我可以选

择 $\delta > 0$, 使得只要 $|x-a| < \delta$, 就有 $|g(x) - g(a)| < \lambda$.

我们要综合考虑. 由于我和 Smiddy 的游戏, 我们知道只要 $|x-a| < \delta$, 就有 $|g(x) - g(a)| < \lambda$. 而你和 Smiddy 的游戏显示, 如果 $|y - g(a)| < \lambda$, 那么 $|f(y) - L| < \varepsilon$. 我们不管 Smiddy, 用 $f(g(a))$ 替换 L, 用 $g(x)$ 替换 y. 可以看到, 只要 $|x-a| < \delta$, 就有 $|f(g(x)) - f(g(a))| < \varepsilon$. 这表示, 如果我直接与你对抗, 我总是可以做一次合情理的移动, 不管 ε 是什么 (只要它为正). 因此, 我们实际上就证明了

$$\lim_{x\to a} f(g(x)) = f(g(a)),$$

其中 g 在 $x = a$ 上连续且 f 在 $g(a)$ 上连续. 当然, 如果 f 和 g 都处处连续, 那么复合函数 $f \circ g$ 也处处连续.

我们可以对论证进行修正, 以便包括 $x \to \infty$ 或 $x \to -\infty$ 而不是 $x = a$ 的情况. 由于右边不能是 $g(\infty)$, 故我们必须对陈述稍作修改. 最好的做法就是

$$\lim_{x\to\infty} f(g(x)) = f\left(\lim_{x\to\infty} g(x)\right).$$



我们也可以对 $x \to -\infty$ 的情况做类似的修改. 我把证明的细节留给你来完成, 但基本思想是你和 Smiddy 的对抗是不变的, 但我和 Smiddy 的对抗会稍有不同：我选取 N 而不是 δ, 且不等式 $|x-a| < \delta$ 必须用 $x > N$ 或 $x < N$ 来替换, 这取决于你所处的情况是 $x \to \infty$ 还是 $x \to -\infty$.

我们现在可以建立极限

$$\lim_{x\to\infty} \sin\left(\frac{1}{x}\right) = 0,$$

它在 3.4 节出现过. 事实上, 如果你设 $f(x) = \sin(x)$ 且 $g(x) = 1/x$, 除了 g 在 $x = 0$ 上不连续外, f 和 g 都是处处连续的. 因为

$$\lim_{x\to\infty} g(x) = \lim_{x\to\infty} \frac{1}{x} = 0,$$

我们可以使用上述公式推出结论

$$\lim_{x\to\infty} \sin\left(\frac{1}{x}\right) = \lim_{x\to\infty} f(g(x)) = f\left(\lim_{x\to\infty} g(x)\right) = f(0) = \sin(0) = 0.$$

更直观的一种表达方式是, 当 $x \to \infty$ 时 $1/x \to 0$, 故当 $x \to \infty$ 时 $\sin(1/x) \to \sin(0) = 0$.

A.4.2 介值定理的证明

在 5.1.4 节中, 我们见过介值定理, 它表明如果 f 在 $[a, b]$ 上连续, 且 $f(a) < 0$ 及 $f(b) > 0$, 那么存在某个数 c 使得 $f(c) = 0$. 现在, 我们来看看证明此定理的基本思想.

我们考虑区间 $[a, b]$ 上使得 $f(x) < 0$ 的 x 值的集合. 我们知道 a 在这个集合中, 因为 $f(a) < 0$; 而 b 不在这个集合中. 我们想要求出此集合中最大的数 c, 但这

或许不太可能. 例如, 小于 0 的最大数是什么呢? 没有. 对于任意的负数, 你总是可以找到一个接近 0 的负数, 例如, 将你的数除以 2. 另一方面, 我们可以找到此集合中右边穿插的一个数 c. 特别地, 我们可以坚持说此集合中没有哪个元素在 c 的右边, 而且任意带有端点 c 的开区间至少包括此集合中的一个元素. (这来自于实轴的一个很好的性质 —— 完备性.) 以下是我们知道的, 用符号表示:

(1) 对于任意的 $x > c$, 我们有 $f(x) \geqslant 0$;

(2) 对于任意的区间 $(c-\delta, c)$, 其中 $\delta > 0$, 区间内至少存在一点 x 使得 $f(x) < 0$.

现在该忙起来了. 以下就是重要的问题: $f(c)$ 是什么? 我们假设它是负的. 在这种情况下, 由于 $f(b) > 0$, 故 $c \neq b$. 因为 f 是连续的, 所以当 x 在 c 的附近时, $f(x)$ 的值应该在 $f(c)$ 的附近; 但当 x 在 c 的右边一点点时就会有问题, 因为 $f(x)$ 预期应该是正的, 而 $f(c)$ 为负. 更正式地, 你可以选择 $\varepsilon = -f(c)/2$ (它是正的), 那么你的容忍区间就是 $(3f(c)/2,\ f(c)/2)$, 它仅由负数组成. 我不能选取任何位于 $[a,\ b]$ 中形如 $(c-\delta,\ c+\delta)$ 的区间, 因为任何这样的区间都包含一个大于 c 的 x. 根据上面的条件 (1), 我们知道 $f(x)$ 一定为正, 这表示它不会位于你的容忍区间. 因此, 不可能有 $f(c) < 0$. 直观上, 如果有 $f(c) < 0$, 那么你的穿插仍然有数在它的右边!

或许 $f(c) > 0$. 在这种情况下, 我们不可能有 $c = a$, 因为 $f(a) < 0$. 现在, 当 x 在 c 的附近时, $f(x)$ 的值应该在 $f(c)$ 附近; 特别地, 它们应该是正的. 由于上面的条件 (2), 所以这是个问题. 更明确些, 这一次你可以选择 $\varepsilon = f(c)/2$, 则你的容忍区间是 $(f(c)/2,\ 3f(c)/2)$. 我需要尝试找到一个在 $[a,\ b]$ 中的区间 $(c-\delta,\ c+\delta)$, 使得对于我的区间中的任意 x, $f(x)$ 总是位于你的容忍区间里. 特别地, $f(x) > 0$. 这意味着, 对于 $(c-\delta,\ c)$ 中的所有 x 有 $f(x) > 0$, 这和条件 (2) 是相悖的. 故 $f(c) > 0$ 也不可能. 如果它是真的, 那么我们可以将穿插再向左边挪一些, 因此它不会是 c.

剩下的是什么呢? 唯一可能就是 $f(c) = 0$, 因此, 我们证明了该定理. 顺便要说的是, 我们很容易将情况改为 $f(a) > 0$ 及 $f(b) < 0$ 的情况. 你可以稍稍改写一下证明, 或者设 $g(x) = -f(x)$ 并对 g 而不是 f 应用该定理.

A.4.3　最大–最小定理的证明

现在我们来证明 5.1.6 节的最大 – 最小定理. 其基本思想是, 假定我们有一个在闭区间 $[a,\ b]$ 上连续的函数 f, 我们断言, 该区间上存在某个数 c 使 f 达到最大值. 正如我们看到的, 这表示 $f(c)$ 大于或等于其他 $f(x)$ 的值, 其中 x 在整个区间 $[a,\ b]$ 上漫游.

证明如下. 我们想要证明的是, 你可以放置某条水平线 $y = N$, 使得所有的函数值 $f(x)$ 都位于这条线的下方. 如果做不到这一点, 那么函数就会在 $[a,\ b]$ 内的某处变得越来越大, 而不会有最大值. 因此, 我们假设你画不出这样的一条线. 那么,

对于每一个正数 N, 在 $[a, b]$ 中存在某个点 x_N 使得 $f(x_N)$ 在水平线 $y=N$ 的上方即我们找到了若干个点 x_N, 对于每一个 N, 都有 $f(x_N)>N$. 我们在 x 轴上用 X 将它们标出来.

这些标记点在哪里呢？有无穷多个这样的点. 因此, 如果我们将区间 $[a, b]$ 分成两半得到两个新的区间, 它们中的某一个定然包含无穷多个标记点. 它们可能都包含无穷多个标记点, 但不可能都只包含有限个标记点, 否则总的标记点将是有限的. 让我们把注意力集中在原始区间中包含无穷多个标记点的那一半上. 如果它们都如此, 那就选择你最喜欢的那个 (这没有关系的). 现在, 我们用新的更小的区间重复这个练习：将它分成两半. 其中之一一定包含无穷多个标记点. 只要你喜欢, 我们就继续做这个练习, 你会得到一个变得越来越小的区间的集合, 一个套一个, 并且每一个都包含无穷多个标记点. 我们将这些区间一个一个地堆在一起, 如图 A-9 所示.

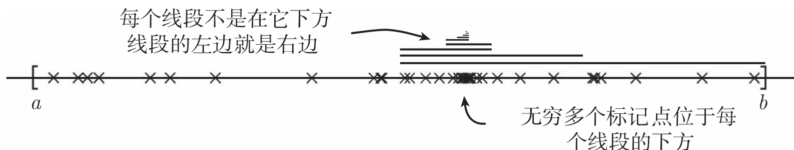


图 A-9

直观上, 必有实数存在于所有这些区间之中,[①] 我们称之为数 q. $f(q)$ 是什么呢？我们可以使用 f 的连续性来获得一些信息. 事实上, 我们知道

$$\lim_{x\to q} f(x)=f(q).$$

因此, 打个比方, 如果你选取的 ε 是 1, 那么我应该能够找到一个区间 $(q-\delta,\ q+\delta)$, 使得对于所有该区间中的 x 都有 $|f(x)-f(q)|<1$. 问题是, 这个区间 $(q-\delta,\ q+\delta)$ 包含了无穷多个标记点! 因为不管 δ 多么小, 我们选择的最后一个小区间都会位于 $(q-\delta,\ q+\delta)$ 内. 这才是问题所在：所有这些标记点都应该在区间 $(q-\delta, q+\delta)$ 内, 当你对其中任意一个点取 f 值时, 会得到一个介于 $f(q)-1$ 和 $f(q)+1$ 之间的数. 因此, 不管 $f(q)$ 是什么, 我们都会陷入困境：某些标记点的函数值会远远大于 $f(q)+1$. 一切都将失去控制. 因此, 画不出让整个函数位于其下方的直线 $y=N$, 这一假设是错的.

事情还没有结束. 我们有了这条线 $y=N$, 它位于 $y=f(x)$ 在 $[a, b]$ 的图像的上方, 现在, 我们需要将它向下移动, 直到它接触到该图像以便求最大值. 因此, 我们选取尽可能小的 N, 使得对于 $[a, b]$ 内的所有 x 有 $f(x)\leqslant N$. (我们再次使用了完备性.) 现在我们需要证明, 对于某个 c 有 $N=f(c)$. 为了求证, 我们要重复在标记点中所使用的技巧, 只是这一次将它们用圈标记出来. 我们选取一个正整数 n,

① 同样, 我们需要使用实轴的完备性来证明. 事实上, 一定只存在一个这样的数 —— 你知道为什么吗?

在 $[a,\ b]$ 中一定能够找到某个数 c_n, 使得 $f(c_n) > N - 1/n$. 如若不然, 我们就应该在 $y = N - 1/n$(或更低处) 而不是 $y = N$ 处画那条线. 因此, 存在这样的一个 c_n, 且对于每一个正整数 n 都存在. 我们将这些点圈起来. 有无穷多个这样的点, 当你对它们取 f 值时, 其结果会越来越接近 —— 事实上是任意地接近 ——N. (没有一个值会超过 N, 因为对于所有的 x 有 $f(x) \leqslant N$!) 现在, 我们所要做的就是持续将区间 $[a,\ b]$ 进行二分, 使得每一个小区间都包含无穷多个圈起来的点. 和前面一样, 在所有的区间中都存在一个数 c. 这个数又被圈起来的点所环绕着.

$f(c)$ 是什么呢? 它不可能大于 N, 但或许它会小于 N. 我们假设 $f(c) = M$, 其中 $M < N$, 另外设 $\varepsilon = (N - M)/2$. 由于 f 是连续的, 我们实际上需要

$$\lim_{x\to c} f(x) = f(c) = M.$$

你有你的 ε, 因此我需要找到一个区间 $(c-\delta,\ c+\delta)$, 使得对于在我区间内的 x, $f(x)$ 位于 $(M-\varepsilon,\ M+\varepsilon)$ 中. 问题是 $M + \varepsilon = N - \varepsilon$, 且不管我如何选取 $\delta > 0$, 都有无穷多个圈起来的点位于 $(c-\delta,\ c+\delta)$ 中. 它们其中一些的函数值可能位于 $(M-\varepsilon,\ M+\varepsilon)$ 中, 但由于函数值会变得接近 N, 因而大多数不会位于 $(M-\varepsilon, M+\varepsilon)$ 中. 因此, 我不能移动. 唯一解脱的方法就是 $f(c) = N$. 这表示 c 是函数取得最大值的点. 这样我们就完成了求证!

要得到定理的最小值的形式, 只需要将定理重新应用到 $g(x) = -f(x)$ 上就可以了. 毕竟, 如果 c 是 g 取得最大值的点, 那么它就是 f 取得最小值的点.

A.5　再谈指数函数和对数函数

在 9.2 节中, 我们发展了指数函数和对数函数的理论, 最终的发现就是

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^x = \mathrm{e}^x \quad 和 \quad \frac{\mathrm{d}}{\mathrm{d}x}\ln(x) = \frac{1}{x}.$$

当时还有个不精确的结尾：我们断言

$$\lim_{h\to 0^+}(1+h)^{1/h}$$

存在, 并称之为 e, 但我们并没有证明过它. 直接证明上述极限存在是可能的, 但这提供不了任何特别的信息. 反之, 我假设你已经学了积分和微积分基本定理 (见第 16 章和第 17 章), 从而我可以用一个不同的方法解决问题. 事实上, 一切都是从对数函数开始的.

我们先根据规则定义一个函数 F,

$$F(x) = \int_1^x \frac{1}{t}\mathrm{d}t$$

对于所有的 $x > 0$ 成立. 这个函数基于另一个函数的积分, 就这类函数请参见 17.1 节. 现在, 我知道你可以写出

$$F(x)=\int_1^x \frac{1}{t}\mathrm{d}t=\ln|t|\Big|_1^x=\ln|x|-\ln|1|=\ln(x),$$

因为 $x>0$ 且 $\ln(1)=0$. 问题是, 我们的行动过早了! 如果真要以恰当的方式求解, 就不能使用 $\int 1/t\mathrm{d}t=\ln|t|+C$ 这一事实. 实际上, 这是我们想要证明的事情之一. 目前为止, 我们不能假设 $F(x)=\ln|x|$, 那就让我们从证明它开始吧.

让我们写出函数 F 的一些有趣性质. 根据微积分第一基本定理, F 的导数为

$$F'(x)=\frac{\mathrm{d}}{\mathrm{d}x}\int_1^x \frac{1}{t}\mathrm{d}t=\frac{1}{x}.$$

因此, F 可导, 这意味着它是连续的 (见 5.2.11 节). 接下来, 我们设 $x=1$, 从而得

$$F(1)=\int_1^1 \frac{1}{t}\mathrm{d}t=0,$$

因为若积分上下限相等且函数在那里确实有定义, 则任何函数的积分都是 0(见 16.3 节). 极限

$$\lim_{x\to\infty}F(x)$$

如何呢? 事实上, 根据反常积分的定义 (见 20.2 节), 我们有

$$\lim_{x\to\infty}F(x)=\lim_{x\to\infty}\int_1^x \frac{1}{t}\mathrm{d}t=\int_1^\infty \frac{1}{t}\mathrm{d}t=\infty.$$

反常积分 $\int_1^\infty 1/t\mathrm{d}t$ 发散, 我们必须非常小心提到. 最初证明它的发散性时, 我们使用了公式 $\int 1/t\mathrm{d}t=\ln|t|+C$, 但我们现在不能这样做! 而是使用了积分判别法来说 $\int_1^\infty 1/t\mathrm{d}t$ 和 $\sum\limits_{n=1}^{\infty}1/n$ 同时收敛或同时发散; 然后使用 22.4.3 节中的论证来证明该级数发散; 故该积分也发散. 因此我们有

$$F(1)=0 \text{ 和 } \lim_{x\to\infty}F(x)=\infty.$$

由于 F 连续, 介值定理 (见 5.1.4 节) 表明, 一定存在一个数 e 使得 $F(\mathrm{e})=1$. 毕竟, 1 介于 0 和 ∞ 之间! 此外, 对于所有的 $x>0$, $F'(x)=1/x>0$, 我们因此知道 F 总是递增的. 因此, 不可能存在其他的数 c, 使得 $F(c)=1$. 我们已经有了 e 的正式定义:

$$\mathrm{e} \text{ 是唯一使得 } \int_1^{\mathrm{e}} \frac{1}{t}\mathrm{d}t=1 \text{ 的数.}$$

现在, 我们选取一个有理数 α 并定义

$$G(x)=F(x^\alpha)=\int_1^{x^\alpha}\frac{1}{t}\mathrm{d}t.$$

使用 17.5.2 节中描述的变形 2 的技巧, 可以看到

$$G'(x)=\frac{d}{dx}\int_1^{x^\alpha}\frac{1}{t}\mathrm{d}t=\alpha x^{\alpha-1}\frac{1}{x^\alpha}=\alpha\cdot\frac{1}{x}.$$



(不使用对数函数求导, 假设我们知道 $\frac{\mathrm{d}}{\mathrm{d}x}(x^\alpha)=\alpha x^{\alpha-1}$. 如果只知道对于正整数是成立的, 正如 6.1 节所述, 那么我们来看看是否可以对于所有的有理数来证明这个事实.) 另一方面, 我们知道 $F'(x)=1/x$, 因此, 上述方程暗示了 $G'(x)=\alpha F'(x)$. 由于 α 是常数, 我们看到 $G(x)=\alpha F(x)+C$, 其中 C 是常数. 特别地, 如果我们设 $x=1$, 此方程变为 $G(1)=\alpha F(1)+C$. 现在有 $G(1)=F(1^\alpha)=F(1)=0$, 故 $C=0$. 由于 $G(x)=F(x^\alpha)$, 我们就证明了 $F(x^\alpha)=\alpha F(x)$, 对于任意有理数 α 及 $x>0$ 成立. 事实上, 由于 F 连续, 结果对于任意实数 α 一定也适用! 现在我们设 $x=\mathrm{e}$, 会看到 $F(\mathrm{e}^\alpha)=\alpha F(\mathrm{e})=\alpha$, 因为 $F(\mathrm{e})=1$. 我们将 α 变为 x, 这样就证明了 $F(\mathrm{e}^x)=x$. 因此, F 是 e^x 的反函数, 这表示 $F(x)=\ln(x)$. 因为我们知道 $F'(x)=1/x$, 这就证明了 $\frac{\mathrm{d}}{\mathrm{d}x}\ln(x)=1/x$. 现在, 如果 $y=\mathrm{e}^x$, 那么 $x=\ln(y)$, 故

$$\frac{\mathrm{d}x}{\mathrm{d}y}=\frac{1}{y}=\frac{1}{\mathrm{e}^x};$$

根据链式法则, $\mathrm{d}y/\mathrm{d}x=\mathrm{e}^x$. 因此, 我们对 $\ln(x)$ 和 e^x 求了导且证明了 e 存在!

现在, 我们要做的就是证明

$$\lim_{h\to0^+}(1+h)^{1/h}=\mathrm{e}.$$

这十分简单：令 $y=(1+h)^{1/h}$, 于是 $\ln(y)=\ln(1+h)/h$. 故根据 9.4.3 节中使用的论证 (或洛必达法则)

$$\lim_{h\to0^+}\ln(y)=\lim_{h\to0^+}\frac{\ln(1+h)}{h}=1.$$

当然, 如果当 $h\to0^+$ 时 $\ln(y)\to1$, 那么当 $h\to0^+$ 时 $y\to\mathrm{e}^1=\mathrm{e}$. 这就证明了上述极限. 关键是, 一旦你知道 $\ln(x)$ 关于 x 的导数是 $1/x$, 那么其余的一切对你来说就很容易了.

A.6 微分与极限

在这一节, 我们将证明一些涉及微分和极限的结论. 更确切地说, 我们要处理函数的常数倍、函数的和与差的求导, 以及乘积法则、商法则与链式求导法则. 然后, 我们将证明极值定理、罗尔定理、中值定理以及线性化中的误差公式. 最后, 我们会看到分段函数的导数以及洛必达法则的证明.

A.6.1 函数的常数倍

假设 y 是关于 x 的一个可导函数, c 是某个常数. 我们想要证明

$$\frac{\mathrm{d}}{\mathrm{d}x}(cy)=c\frac{\mathrm{d}y}{\mathrm{d}x}.$$

这相当简单. 我们用 $y=f(x)$ 定义 f, 那么上述方程的左边就是

$$\lim_{\Delta x\to0}\frac{cf(x+\Delta x)-cf(x)}{\Delta x}.$$

你所要做的就是从分子中提取一个 c 的因子并将它拖到极限之外. 这是在 A.2.2 节结尾部分证明过的：

$$\lim_{\Delta x\to 0}\frac{cf(x+\Delta x)-cf(x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{c(f(x+\Delta x)-f(x))}{\Delta x}$$
$$=c\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

右边正好是 $cf'(x)$, 它和 $c\,(\mathrm{d}y/\mathrm{d}x)$ 一样. 这样我们就完成了求证.

A.6.2 函数的和与差

假设 u 和 v 都是 x 的可导函数, 我们想要证明的是

$$\frac{\mathrm{d}}{\mathrm{d}x}(u+v)=\frac{\mathrm{d}u}{\mathrm{d}x}+\frac{\mathrm{d}v}{\mathrm{d}x},$$

以及类似地用减号代替加号. 这几乎没什么可证的. 记 $u=f(x)$ 及 $v=g(x)$, 那么上述方程的左边就是

$$\lim_{\Delta x\to 0}\frac{f(x+\Delta x)+g(x+\Delta x)-(f(x)+g(x))}{\Delta x}.$$

你所要做的就是重新整理这个和, 并拆分极限, 这是在 A.2.1 节中证明过的, 上述极限等于

$$\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}+\lim_{\Delta x\to 0}\frac{g(x+\Delta x)-g(x)}{\Delta x}.$$

这就是 $f'(x)+g'(x)$, 它等于我们想要证明的方程的右边. 用减号替换加号的情况也同样简单!

A.6.3 乘积法则的证明

对于乘积法则和商法则的证明, 我们将继续使用记号 $\mathrm{d}y/\mathrm{d}x$ 而不是 $f'(x)$, 因为使用前者更易于理解概念. 正如 5.2.7 节所述, 我们有

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x},$$

其中 Δy 是将 x 变为 $x+\Delta x$ 时 y 的变化量.

因此, 我们想要证明的乘积法则说的就是

$$\frac{\mathrm{d}}{\mathrm{d}x}(uv)=v\frac{\mathrm{d}u}{\mathrm{d}x}+u\frac{\mathrm{d}v}{\mathrm{d}x}.$$

假设我们将 x 变为 $x+\Delta x$, 那么 u 就变为 $u+\Delta u$, v 就变为 $v+\Delta v$. 而 uv 就变为 $(u+\Delta u)(v+\Delta v)$. 这个变化量有多大呢？我们取原来的量与新的量的差来看看：

$$\Delta(uv)=(u+\Delta u)(v+\Delta v)-uv.$$

展开并化简, 得到

$$\Delta(uv)=v\Delta u+u\Delta v+\Delta u\Delta v.$$

现在用该式除以 Δx. 对于最后一项, 我们要多除以一个 Δx, 再乘以这个量使方程两边保持平衡. 结果是

$$\frac{\Delta(uv)}{\Delta x}=v\frac{\Delta u}{\Delta x}+u\frac{\Delta v}{\Delta x}+\frac{\Delta u}{\Delta x}\frac{\Delta v}{\Delta x}\Delta x.$$

如果你取当 $\Delta x\to 0$ 时的极限, 那么所有的比率都会趋于相应的导数, 但最后一个 Δx 的因子会趋于 0, 即



$$\frac{\mathrm{d}}{\mathrm{d}x}(uv)=v\frac{\mathrm{d}u}{\mathrm{d}x}+u\frac{\mathrm{d}v}{\mathrm{d}x}+\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x}\times 0.$$

由于最后一项为 0, 我们就证明了乘积法则. 现在, 你应该尝试写出一个使用 $f(x)$ 记号 (形式 1) 的证明了.

A.6.4 商法则的证明

现在我们想要证明

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{u}{v}\right)=\frac{v\frac{\mathrm{d}u}{\mathrm{d}x}-u\frac{\mathrm{d}v}{\mathrm{d}x}}{v^2}.$$

同样, 当 x 变为 $x+\Delta x$ 时, 我们知道 u 和 v 就会分别变为 $u+\Delta u$ 及 $v+\Delta v$. 而 u/v 就变为 $(u+\Delta u)/(v+\Delta v)$. 这个变化量是

$$\Delta\left(\frac{u}{v}\right)=\frac{u+\Delta u}{v+\Delta v}-\frac{u}{v}.$$

我们对上式通分并消除 $uv-uv$, 得到

$$\Delta\left(\frac{u}{v}\right)=\frac{v\Delta u-u\Delta v}{v^2+v\Delta v}.$$

将上式除以 Δx, 再用 Δx 和分母中的 Δv 的项相乘并相除, 得到

$$\frac{\Delta\left(\frac{u}{v}\right)}{\Delta x}=\frac{v\frac{\Delta u}{\Delta x}-u\frac{\Delta v}{\Delta x}}{v^2+v\frac{\Delta v}{\Delta x}\Delta x}.$$

现在令 $\Delta x\to 0$. 所有的分式都变为导数, 并且分母中的最后一个因子趋于 0, 因此我们得到结果

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{u}{v}\right)=\frac{v\frac{\mathrm{d}u}{\mathrm{d}x}-u\frac{\mathrm{d}v}{\mathrm{d}x}}{v^2+v\frac{\mathrm{d}v}{\mathrm{d}x}\times 0}.$$

由于分母中的最后一项是 0, 我们证明了商法则.

A.6.5 链式求导法则的证明

假设 y 是 u 的可导函数, 而 u 本身是 x 的可导函数. 我们想要证明

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}.$$

第一眼看上去这也没什么, 可使用 Δ 记号写出

$$\frac{\Delta y}{\Delta x}=\frac{\Delta y}{\Delta u}\frac{\Delta u}{\Delta x}$$

并取极限. 不幸的是, Δu 有时可能为 0, 而这会导致整个等式无效. 因此, 我们使用函数记号. 令 f 和 g 都是可导的, 并设 $h(x)=f(g(x))$. 我们想要证明

$$h'(x)=f'(g(x))g'(x).$$

如果 g 在 x 附近是常数, 那么 h 也是, 因此等式两边都是 0. 否则, 我们知道

$$h'(x) = \lim_{\Delta x \to 0} \frac{h(x+\Delta x) - h(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{\Delta x}.$$

用该分式乘以并除以 $g(x+\Delta x) - g(x)$, 对于无穷多个在 0 附近的 Δx 值, 这一定是非零的, 然后我们将极限拆分得到

$$h'(x) = \lim_{\Delta x \to 0} \frac{f(g(x+\Delta x)) - f(g(x))}{g(x+\Delta x) - g(x)} \times \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x}.$$

右边的极限就是 $g'(x)$, 但左边是什么呢? 求解技巧是设 $\varepsilon = g(x+\Delta x) - g(x)$. 那么, 左边极限的分子中的量 $g(x+\Delta x)$ 可以被写作 $g(x)+\varepsilon$, (你知道这是为什么吗?) 而分母正是 ε 本身. 因此我们有

$$h'(x) = \lim_{\Delta x \to 0} \frac{f(g(x)+\varepsilon) - f(g(x))}{\varepsilon} \times g'(x).$$

现在, 当 $\Delta x \to 0$ 时, ε 会怎样呢? 由于 g 可导, 由 5.2.11 节可知 g 连续. 特别地, 有

$$\lim_{\Delta x \to 0} g(x+\Delta x) = g(x).$$

如果从两边减去 $g(x)$, 那么你会看到, 当 $\Delta x \to 0$ 时 $\varepsilon \to 0$. 这表示, 在 $h'(x)$ 的表达式中, 我们可以用 $\varepsilon \to 0$ 替换 $\Delta x \to 0$, 得到

$$h'(x) = \lim_{\varepsilon \to 0} \frac{f(g(x)+\varepsilon) - f(g(x))}{\varepsilon} \times g'(x).$$

第一项正是 $f'(g(x))$, 故 $h'(x) = f'(g(x))\, g(x)$. 这样, 我们就证明了链式求导法则.

A.6.6 极值定理的证明

在 11.1.2 节中, 我们陈述了极值定理. 它说的是, 如果 f 在 $x=c$ 有一个局部最大值或局部最小值, 那么 $x=c$ 是 f 的一个临界点. 这表示, 或者 $f'(c)$ 不存在, 或者 $f'(c)=0$.

为了证明这一点, 我们首先假设 f 在 $x=c$ 有一个局部最小值. 如果 $f'(c)$ 不存在, 那么它就是一个临界点, 这正是我们所希望的. 另一方面, 如果 $f'(c)$ 存在, 那么

$$f'(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}.$$

由于 f 在 c 上有一个局部最小值, 因而我们知道当 $c+h$ 非常接近 c 时, $f(c+h) \geqslant f(c)$. 当然, 只有当 h 接近于 0 时, $c+h$ 才会非常接近 c. 对于这样的 h, 上述分式中的分子 $f(c+h) - f(c)$ 一定是非负的. 当 $h>0$ 时, 量

$$\frac{f(c+h) - f(c)}{h}$$

是正的 (或 0); 但是当 $h<0$ 时, 此量是负的 (或 0). 因此右极限

$$\lim_{x \to c^+} \frac{f(c+h) - f(c)}{h}$$

一定大于或等于 0, 而同样的左极限是小于或等于 0. 由于双侧极限存在, 故左极限等于右极限; 唯一的可能性就是它们都是 0. 这就证明了 $f'(c)=0$, 故 $x=c$ 是 f 的一个临界点.



如果 f 在 $x=c$ 有一个局部最大值会如何呢? 我把这个论证过程留给你来完成. 唯一的区别就是, 当 h 接近于 0 时, 量 $f(c+h)-f(c)$ 是负的 (或 0).

A.6.7 罗尔定理的证明

假设 f 在 $[a, b]$ 上连续, 在 (a, b) 内可导, 且满足条件 $f(a)=f(b)$. 接下来, 我们想要证明在 (a, b) 内存在一个数 c, 使得 $f'(c)=0$. 为了求证, 我们使用最大最小值定理来说明 f 在 $[a, b]$ 上有一个全局最大值和一个全局最小值. 如果最大值或最小值中任一个出现在 (a, b) 内的某个数 c 上, 那么极值定理告诉我们 $f'(c)=0$. (我们知道 $f'(c)$ 存在, 因为 f 在 (a, b) 内可导.) 其他的唯一可能性就是全局最大值和全局最小值都出现在端点 a 和 b 上. 在这种情况下, 由于 $f(a)=f(b)$, 该函数一定为常数, 因此, (a, b) 内的每一个数 c 都满足 $f'(c)=0$. 这就是完整的证明!

A.6.8 中值定理的证明

现在, 我们知 f 在 $[a, b]$ 上连续, 在 (a,b) 内可导, 但我们不假设 $f(a)=f(b)$. 中值定理表明, 在 (a, b) 内存在某个 c 满足

$$f'(c)=\frac{f(b)-f(a)}{b-a}.$$

为了证明这一点, 我们定义一个新的函数 g:

$$g(x)=f(x)-\frac{f(b)-f(a)}{b-a}(x-a).$$

它看起来有点复杂, 但实际上我们只是从 $f(x)$ 中减去了线性函数 $x-a$ 的一个常数倍, 并称之为 g. 因此, 函数 g 也在 $[a, b]$ 上连续且 (a, b) 内可导, 且可知

$$g(a)=f(a)-\frac{f(b)-f(a)}{b-a}(a-a)=f(a),$$

$$g(b)=f(b)-\frac{f(b)-f(a)}{b-a}(b-a)=f(a).$$

因此, 我们证明了 $g(a)=g(b)$, 这表示我们可以应用罗尔定理了! 结果是, 存在一个数 c 使得 $g'(c)=0$. 现在, 我们只需要对 g 求导来看看这对于 f 意味着什么. 由于量 $f(b)-f(a)$ 和 $b-a$ 都是常数, 我们得到

$$g'(x)=f'(x)-\frac{f(b)-f(a)}{b-a}.$$

现在, 将 $x=c$ 代入. 由于 $g'(c)=0$, 我们有

$$0=f'(c)-\frac{f(b)-f(a)}{b-a}.$$

这表示

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

这正好是我们想要证明的!

A.6.9 线性化的误差

让我们来整理另外一个不精确的结果. 在 13.2 节, 我们看到函数 f 关于 $x = a$ 的线性化 L, 其中 a 是 f 定义域内的某个数:

$$L(x) = f(a) + f'(a)(x - a).$$

如果 x 在 a 的附近, 我们可以使用 $L(x)$ 来估算 $f(x)$ 的值. 我们的错误可能有多大呢? 根据 13.2.4 节的公式, 如果 f'' 在 x 和 a 之间存在, 那么

$$|\text{误差}| = \frac{1}{2}|f''(c)||x - a|^2,$$

这里的 c 是介于 x 和 a 之间的某个数. 我们来证明这个公式. 首先, 我们称误差项为 $r(x)$; 由于 $r(x)$ 是 $f(x)$ 的真值和猜测值的差, 故猜测值就是线性化 $L(x) = f(a) + f'(a)(x - a)$. 我们有

$$r(x) = f(x) - L(x) = f(x) - f(a) - f'(a)(x - a).$$

现在, 聪明的做法是将 x 固定为一个常数并且令 a 为变量. 由此启发得到

$$g(t) = f(x) - f(t) - f'(t)(x - t).$$

因此, 只有 $t = a$ 时, 才有误差 $r(x)$. 即, 误差为 $g(a)$. 注意

$$g(x) = f(x) - f(x) - f'(x)(x - x) = 0.$$

我们求 g 关于 t 的导数. 项 $f(x)$ 是常数, 故其导数为 0. 此外, 我们需要用乘积法则来处理 $f'(t)(x - t)$. 总之, 我们得到

$$g'(t) = 0 - f'(t) - (f'(t) \times (-1) + f''(t)(x - t)) = -f''(t)(x - t).$$

特别地, 我们有

$$g'(x) = -f''(x)(x - x) = 0.$$

目前为止, 我们所做的一切都是非常合理的. 现在, 我们必须做一些看起来有些疯狂的事情. 请记住, 我们想要证明误差是 $\frac{1}{2}f''(c)(x - a)^2$, 其中 c 介于 x 和 a 之间. 由于误差是 $g(a)$, 这就暗示了 $g(t)$ 形如 $K(x - t)^2$, 其中 K 是某个不依赖于 t 而只依赖于 x 和 a 的数. 即使这不完全正确, 但它或许可以解释我们为什么会令

$$h(t) = g(t) - K(x - t)^2.$$

你看, 当对它关于 t 求导时, 保持 x 为常数, 会得到

$$h'(t) = g'(t) + 2K(x - t).$$

这又怎么样? 我们可以使用中值定理 (见 11.3 节) 得到

$$h'(c) = \frac{h(x) - h(a)}{x - a}$$

对于某个介于 x 和 a 之间的 c 成立. 我们可以使用上述等式对 $h'(c)$、$h(x)$ 及 $h(a)$ 做替换:

$$\begin{aligned} g'(c) + 2K(x-c) &= \frac{(g(x) - K(x-x)^2) - (g(a) - K(x-a)^2)}{x-a} \\ &= \frac{-g(a) + K(x-a)^2}{x-a}, \end{aligned}$$

因为 $g(x) = 0$. 因为 $g'(c) = -f''(c)(x-c)$, 最后一个方程可以重新整理为

$$g(a) - K(x-a)^2 = (x-a)(x-c)(f''(c) - 2K).$$

我们的任务快完成了, 但仍然有一个问题. 我们不能处理因子 $(x-c)$, 因为在我们的误差项中没有它! 唯一一种消除它的可能就是左边等于 0, 即应该选取 K 使得 $g(a) - K(x-a)^2 = 0$. 事实上, 如果 $K = g(a)/(x-a)^2$, 那么上述方程变为

$$0 = (x-a)(x-c)\left(f''(c) - \frac{2g(a)}{(x-a)^2}\right).$$

由于 $x \neq a$ 且 $x \neq c$, 我们一定会有

$$f''(c) - \frac{2g(a)}{(x-a)^2} = 0,$$

这表示 $g(a) = \frac{1}{2}f''(c)(x-a)^2$. 由于 $g(a) = r(x)$ 是我们要找的误差, 因此我们完成了证明.

A.6.10 分段函数的导数

假定 f 以分段的形式定义为

$$f(x) = \begin{cases} f_1(x) & \text{若} x > a, \\ f_2(x) & \text{若} x \leqslant a. \end{cases}$$

(你可以将 $x > a$ 改为 $x \geqslant a$, 将 $x \leqslant a$ 改为 $x < a$; 这无关紧要.) 不管怎样, 在 6.6 节中, 我们考虑了一个问题, 就是 f 是否在 a 上可导. 我们假设如果函数 f_1 和 f_2 在 $x = a$ 处互相匹配, 则它们的导数 f_1' 和 f_2' 在 $x = a$ 处也互相匹配, 那么 f 在 a 上可导. 我们如何来证明呢? 首先要注意 f_1 和 f_2 在 $x = a$ 处互相匹配的意思是

$$\lim_{x \to a^+} f_1(x) = \lim_{x \to a^-} f_2(x) = f(a).$$

这就确保了 f 至少是连续的. 现在, 我们还要假设它们的导数也互相匹配, 这意味着 f_1 在最接近 a 的右侧是可导的, f_2 在最接近 a 的左侧是可导的, 以及

$$\lim_{x \to a^+} f_1'(x) = \lim_{x \to a^-} f_2'(x) = L,$$

其中 L 是某个很好的有限数. 因此, 我们来考虑

$$\frac{f(a+h) - f(a)}{h},$$

其中 h 是某个很小的数且 $h \neq 0$. 如果 $h > 0$, 那么我们可以应用中值定理 (见 11.3 节) 得到

$$\frac{f(a+h)-f(a)}{h}=f_1'(c),$$

其中 c 是介于 a 和 $a+h$ 之间的某个数. (这里我们需要 f 在 $[a,\ a+h]$ 上的连续性.) 根据三明治定理, 当 $h\to 0^+$ 时, 数 c 就被夹在 a 和 $a+h$ 之间, 故当 $h\to 0^+$ 时有 $c\to a^+$. 我们现在看到

$$\lim_{h\to 0^+}\frac{f(a+h)-f(a)}{h}=\lim_{h\to 0^+}f_1'(c)=\lim_{c\to a^+}f_1'(c)=L.$$

同理得左极限, 只是我们要使用 f_2' 代替 f_1':

$$\lim_{h\to 0^-}\frac{f(a+h)-f(a)}{h}=\lim_{h\to 0^-}f_2'(c)=\lim_{c\to a^-}f_2'(c)=L.$$

左极限和右极限都等于 L, 因此, 我们证明了 $f'(a)$ 存在且它也等于 L.

A.6.11 洛必达法则的证明

我们来证明洛必达法则 (见第 14 章). 确切地说, 假设我们有两个函数 f 和 g, 它们在某个包含点 a 的区间上可导 (但或许不在 a 本身), 且 $f(a)=g(a)=0$; 此外, 除了可能在 a 处, $g'(x)\neq 0$. 那么, 我们需要证明

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)},$$

假设右边的极限存在. 我们需要一个形式略有不同的中值定理, 它被称为柯西中值定理: 如果 f 和 g 在 $[A,\ B]$ 上连续, 在 $(A,\ B)$ 内可导, 且在 $(A,\ B)$ 上 $g'(x)\neq 0$, 那么, 在 $(A,\ B)$ 内存在某个 C 使得

$$\frac{f'(C)}{g'(C)}=\frac{f(B)-f(A)}{g(B)-g(A)}.$$

我们首先来证明它, 然后用它来证明洛必达法则. 顺便提一句, 请注意, 如果对于所有的 x 有 $g(x)=x$, 那么 $g'(x)=1$, 并且上述方程变为

$$f'(C)=\frac{f(B)-f(A)}{B-A}.$$

这正好是常规的中值定理! 尽管如此, 它对我们没有太多帮助. 让我们回到原始方程中去看看右边的分母吧, 即 $g(B)-g(A)$. 这不可能等于 0; 要是那样的话, 则 $g(B)=g(A)$, 这表示根据罗尔定理 (见 11.2 节), 对于在 $(A,\ B)$ 内的某个 C, $g'(C)=0$. 因此, 右边有意义. 现在, 我们定义一个新的函数 h:

$$h(x)=f(x)-\left(\frac{f(B)-f(A)}{g(B)-g(A)}\right)g(x)$$

对于在 $(A,\ B)$ 内的所有的 x 成立. (将这个函数与 A.6.8 节中的常规中值定理的证明中的函数 g 做比较.) 不管怎样, 我们来写出有关这个函数的某些事实吧. 首先, 计算 $h(A)$ 和 $h(B)$. 我们有

$$\begin{aligned}h(A) &= f(A) - \left(\frac{f(B)-f(A)}{g(B)-g(A)}\right)g(A)\\ &= \frac{f(A)g(B)-f(A)g(A)-f(B)g(A)+f(A)g(A)}{g(B)-g(A)}\\ &= \frac{f(A)g(B)-f(B)g(A)}{g(B)-g(A)},\end{aligned}$$

而

$$\begin{aligned}h(B) &= f(B) - \left(\frac{f(B)-f(A)}{g(B)-g(A)}\right)g(B)\\ &= \frac{f(B)g(B)-f(B)g(A)-f(B)g(B)+f(A)g(B)}{g(B)-g(A)}\\ &= \frac{f(A)g(B)-f(B)g(A)}{g(B)-g(A)}.\end{aligned}$$

故 $h(A)=h(B)$. 此外, 注意 h 是可导的, 并且由于 A 和 B 都是常数, 我们有

$$h'(x) = f'(x) - \left(\frac{f(B)-f(A)}{g(B)-g(A)}\right)g'(x).$$

由于 $h(A)=h(B)$, 我们可以使用罗尔定理来推出结论：在 $(A,\ B)$ 内存在一个数 C, 使得 $h'(C)=0$. 这意味着

$$h'(C) = f'(C) - \left(\frac{f(B)-f(A)}{g(B)-g(A)}\right)g'(C) = 0.$$

如果你重新整理这个方程, 就会得到我们想要的结果

$$\frac{f'(C)}{g'(C)} = \frac{f(B)-f(A)}{g(B)-g(A)}.$$

现在, 我们来证明洛必达法则. 由于 $f(a)=g(a)=0$, 我们有

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f(x)-f(a)}{g(x)-g(a)}.$$

如果 $x>a$, 那么在区间 $[a,\ x]$ 上, 我们可以使用柯西中值定理 (就是我们刚刚证明的) 来说明

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f(x)-f(a)}{g(x)-g(a)} = \lim_{x\to a}\frac{f'(c)}{g'(c)}$$

对于在 $(a,\ x)$ 内的某个 c 成立. 否则, 如果 $x<a$, 那么我们会有相同的结果, 只是 c 在 $(x,\ a)$ 中. (注意, 我们使用的事实是, 除了可能在 a 外, g' 不为 0; 这是柯西中值定理的一个条件.) 当然, 数 c 依赖于 x 的值; 而我们看到, 当 $x\to a$ 时, 也会有 $c\to a$. 因此, 我们有

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(c)}{g'(c)} = \lim_{c\to a}\frac{f'(c)}{g'(c)}.$$

所剩的工作就是将 c 看成虚拟变量并将它改为 x, 这样, 我们就完成了洛必达法则的证明.

嗯, 其实证明不算完整. 我们还没有证明 ∞/∞ 的情况, 也没有证明 $x \to \infty$ (或 $-\infty$) 时的情况. 如果你敢于挑战, 就请尝试将上述证明应用到这些情况中吧, 这是一个极棒的练习.



A.7 泰勒近似定理的证明



现在, 我们来看看如何证明 24.1.3 节中的泰勒近似定理吧. 该定理说的是：如果 f 在 $x = a$ 处是光滑的, 那么, 在所有 N 次或 N 次以下的多项式中, 对于在 a 附近的 x 的 $f(x)$ 的最佳近似就是 N 阶泰勒多项式 P_N 它由下式给出：

$$P_N(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(N)}(a)}{N!}(x-a)^N.$$

我们的计划是, 证明该定理是如何从 24.1.4 节的完整泰勒定理中推导出来的. 我省略了完整泰勒定理的证明, 因为从大多数教科书或在搜索引擎上输入 “泰勒定理的证明” 能找到. 你不容易找到的是这个近似定理的证明, 因此我们就来看看它吧.

首先, 让我们设 $a = 0$ 来简化这个问题. 由于我们假设完整泰勒定理已经被证明了, 因此可知 $f(x) = P_N(x) + R_N(x)$, 其中

$$P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(0)}{n!} x^n$$

是一个 N 次多项式, 也知

$$R_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!} x^{N+1},$$

其中 c 介于 0 和 x 之间. (请记住, 我们设了 $a = 0$, 因此, 形如 $(x-a)^n$ 的因子就变为 x^n, 形如 $f^{(n)}(a)$ 的量就变为 $f^{(n)}(0)$.) 我们想要证明的就是：

在所有 N 次或 N 次以下的多项式中, P_N 是在 0 附近 f 的最佳近似.

到底如何证明类似的陈述呢? 上下文中的 “最佳” 又意味着什么呢? 求解技巧是, 另外选取一个次数不超过 N 的多项式, 我们称之为 Q. 由于 Q 不同于 P_N, 所以 Q 至少有一个系数不同于 P_N 中的相应系数. 我们想要证明 $P_N(x)$ 比 $Q(x)$ 更接近 $f(x)$, 至少当 x 接近 0 时如此. 为了看到这两个量有多么接近, 你需查看一下这两个量的差. 因此, 我们真正想要证明的就是不等式 $|f(x) - P_N(x)| < |f(x) - Q(x)|$, 此时取 x 接近 0. 如果这是正确的, 那么就可以推出结论 —— $P_N(x)$ 确实比 $Q(x)$ 更接近理想值 $f(x)$.

为了得到这个不等式, 我们来分别看看两边的情况. 左边是 $f(x) - P_N(x)$ 的绝对值, 这实际上就是余项 R_N. 我们已经有一个 R_N 的表达式, 它包括三个因子, 即 $f^{(N+1)}(c)$、x^{N+1} 及 $1/(N+1)!$. 我们知道 c 介于 0 和 x 之间, 当 $x \to 0$ 时, 根

据三明治定理一定有 $c \to 0$. 由于我们假定 f 非常光滑, 函数 $f^{(N+1)}$ 是连续的. 因此, 当 $x \to 0$ 时有 $c \to 0$, 故得出 $f^{(N+1)}(c) \sim f^{(N+1)}(0)$. 将这三个因子写在一起并取绝对值, 我们有

$$|f(x) - P_N(x)| = |R_N(x)| = \left|\frac{f^{(N+1)}(c)}{(N+1)!}x^{N+1}\right| \sim \frac{|f^{(N+1)}(0)|}{(N+1)!}|x|^{N+1},$$

其中 $x \to 0$. 事实上, 我们可以令 $C = f^{(N+1)}(0)\,/\,(N+1)!$, 要注意 C 只是某个不依赖于 x 的常数. 因此, 我们有

$$|f(x) - P_N(x)| \sim |C||x|^{N+1}, \qquad \text{当 } x \to 0 \text{ 时}.$$

这太棒了. 现在, 我们来看看要证明的不等式的右边. 这个量是 $|f(x) - Q(x)|$. 我们写出 $f(x) = P_N(x) + R_N(x)$, 从而

$$|f(x) - Q(x)| = |P_N(x) + R_N(x) - Q(x)| = |S(x) + R_N(x)|,$$

其中, 我们通过设 $S(x) = P_N(x) - Q(x)$ 将 $P_N(x)$ 和 $Q(x)$ 放在一起. 让我们来好好看看 S. 它是两个次数不超过 N 的不同多项式的差. 因此, S 是一个次数小于或等于 N 的多项式, 但它不是零多项式. 我们假设, 如果用 x 的幂来写 $S(x)$, 它就好像 $S(x) = a_m x^m + \cdots$, 其中 $a_m x^m$ 是最低次数项. 数 m 必然介于 0 和 N 之间, 因为 S 的次数小于或等于 N. 我们知道 S 的行为很像它的最低次数项的行为 (见 21.4.1 节). 即, 当 $x \to 0$ 时, $S(x) \sim a_m x^m$. 另一方面, 我们需要看看 $S(x) + R_N(x)$, 因为这是我们想要的不等式的右边. 我们已经看到了, 当 $x \to 0$ 时 $R_N(x) \sim Cx^{N+1}$, 故 $S(x) + R_N(x)$ 中的最低次数项的行为仍会像 $a_m x^m$ 一样 (请记住, $m \leqslant N$, 故 x^m 是一个次数低于 x^{N+1} 的项). 综述, 我们有

$$|f(x) - Q(x)| = |S(x) + R_N(x)| \sim |a_m||x^m|, \quad \text{当 } x \to 0 \text{ 时}.$$

太棒了! 我们想要证明不等式

$$|f(x) - P_N(x)| < |f(x) - Q(x)|$$

当 x 接近 0 时是成立的. 我们知道, 当 $x \to 0$ 时, $|f(x) - P_N(x)| \sim |C|\,|x|^{N+1}$ 及 $|f(x) - Q(x)| \sim |a_m|\,|x|^m$. 由于 $m < N+1$(及 $|C|$ 与 $|a_m|$ 都是常数), 易知当 x 很小时, $|C|\,|x|^{N+1}$ 比 $|a_m|\,|x|^m$ 小得多. 事实上, 这两个量的比率是

$$\frac{|C||x|^{N+1}}{|a_m||x|^m} = C_1|x|^{N+1-m},$$

其中 $C_1 = |C|\,/\,|a_m|$ 只是另一个常数. 当 $x \to 0$ 时, 右边的量趋于 0. 因此, 当 x 接近 0 时, 以上不等式实际上是成立的. 最终我们完成了泰勒近似定理的证明!

事实上, 有一点我们没有考虑: 假设 $a = 0$. 为了由此推出一般情况, 你只需在上述证明过程中每一处都用被平移的量 $(x - a)$ 替换量 x. 你只需要注意, $(x - a) \to 0$ 和 $x \to a$ 是同一个意思. 我把证明细节留给你来完成. 如果你能通过上述证明做到这点, 那你就太棒了.



附录 B　估 算 积 分

看到定积分时, 我们习惯于通过反导数以及微积分第二基本定理来给出一个确切的答案. 可实际上, 求解一个有用的反导数可能会很困难或者根本不可能. 有时候, 最好的选择是求出一个积分值的近似. 因此, 我们将讨论估算定积分的三种技巧, 以下就是最后这个附录的内容:

- 使用条纹、梯形法则及辛普森法则估算定积分;
- 估算上述近似中的误差.

B.1　使用条纹估算积分

以下是一个完全合理的定积分:

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x.$$

它相当于由 x 轴、曲线 $y=\mathrm{e}^{-x^2}$ 以及直线 $x=0$ 与 $x=2$ 所围成区域的面积, 如图 B-1 所示.

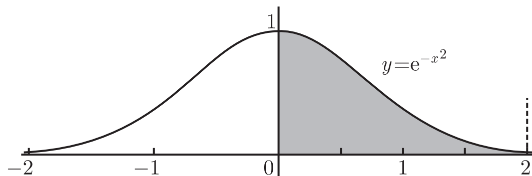


图　B-1

求这样的区域面积或许看起来偏于技术性, 但它有非常大的实际意义. 上述曲线通常被认为是钟形曲线,[①] 而且它是概率论学习的基础. 因此, 特别烦扰的是, 没有简单的好方法来写出反导数

$$\int \mathrm{e}^{-x^2}\mathrm{d}x.$$

实际上, 你可以使用麦克劳林级数把这个积分写成一个无穷级数, 但这也不是简单的好方法. 当前的严峻现实是, 无法将本节最开始的那个定积分的确切值以简洁的方式写出来. (在 16.5.1 节中, 我们已经讨论了这一点.)

另一方面, 我们可以使用黎曼积分的定义求出这个积分的近似值, 即一个估算. 实际上, 在 16.2 节, 我们讨论了划分、区间以及黎曼和. 由于积分是黎曼和的极限,

① 技术上说, 钟形曲线 (或正态分布) 实际上是由方程 $y=\mathrm{e}^{-x^2/2}/\sqrt{2\pi}$ 给出的.

不取极限, 我们就可以得到一个近似. 因此, 为了估算积分

$$\int_a^b f(x)\mathrm{d}x,$$

可以将区间 $[a,\ b]$ 做一个形如

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

的划分, 然后在 $[x_0,\ x_1]$ 中选取一点 c_1, 在 $[x_1,\ x_2]$ 中选取一点 c_2, 以此类推直到在 $[x_{n-1},\ x_n]$ 中选取一点 c_n. 那时, 就可以写出

$$\int_a^b f(x)\mathrm{d}x \approx \sum_{j=1}^{n} f(c_j)(x_j - x_{j-1}).$$

这就是说, 积分近似等于它的一个黎曼和.

所有这一切看起来都很抽象. 我们来看看它在上例中是如何起作用的吧. 我们要从 0 到 2 积分, 因此需要区间 $[0,\ 2]$ 上的一个划分. 该区间上最简单的划分就是这个区间 $[0,\ 2]$, 这相当于选择 $n = 1$、$x_0 = 0$ 及 $x_1 = 2$. 我们只需要在 $[0,\ 2]$ 内选取 c_1. 求出的近似很大程度上依赖于这个选取! 例如, 如果选取 $c_1 = 0$、$c_1 = 1$ 或 $c_1 = 2$, 那么近似就会分别对应图 B-2 所示区域的面积.

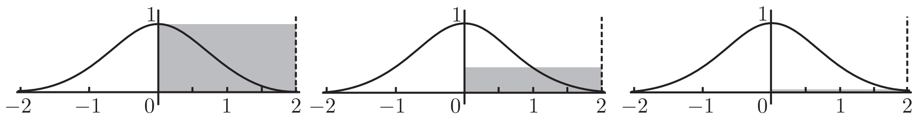


图　B-2

很明显, 第一个估算过高了, 而第三个则估算过低了. 中间的那个不算太糟, 但它仍不完美. 为了计算这三个估算值, 我们使用公式

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx \sum_{j=1}^{n} f(c_j)(x_j - x_{j-1}).$$

我们用 1 替换 n, $\mathrm{e}^{-c_1^2}$ 替换 $f(c_1)$, 0 替换 x_0, 并用 2 替换 x_1, 得到

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx \mathrm{e}^{-c_1^2}(2-0) = 2\mathrm{e}^{-c_1^2}.$$

当 c_1 是 0、1 或 2 时, 这些值分别是 2、$2/\mathrm{e} \approx 0.736$ 及 $2/\mathrm{e}^4 \approx 0.037$. 正如你看到的, 这三个估算有很大的差别!

现在我们来看看, 使用更多的条纹是否可以做得更好. 假设我们取了 $[0,\ 2]$ 上的一个五条划分

$$0 < \frac{1}{2} < 1 < \frac{5}{4} < \frac{3}{2} < 2.$$

因此, $n = 5$, $x_0 = 0$, $x_1 = \frac{1}{2}$, $x_2 = 1$, $x_3 = \frac{5}{4}$, $x_4 = \frac{3}{2}$, $x_5 = 2$. 假设我们选取的数 c_j 是每一个小区间的左端点, 这就表示 $c_1 = 0$, $c_2 = \frac{1}{2}$, $c_3 = 1$, $c_4 = \frac{5}{4}$, $c_5 = \frac{3}{2}$. 将这些

数代入上述近似公式中, 可得

$$\begin{aligned}\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x &\approx \sum_{j=1}^{n} f(c_j)(x_j - x_{j-1})\\&= \sum_{j=1}^{5} \mathrm{e}^{-c_j^2}(x_j - x_{j-1})\\&= \mathrm{e}^{-0^2}\left(\frac{1}{2}-0\right) + \mathrm{e}^{-(1/2)^2}\left(1-\frac{1}{2}\right) + \mathrm{e}^{-1^2}\left(\frac{5}{4}-1\right)\\&\quad + \mathrm{e}^{-(5/4)^2}\left(\frac{3}{2}-\frac{5}{4}\right) + \mathrm{e}^{-(3/2)^2}(2-\frac{3}{2}).\end{aligned}$$



如果你喜欢, 可以再做一些简化, 或者使用计算器或计算机得出其近似到小数点后四位的结果 1.0865. 现在, 你的任务是, 求使用每一个小区间的右端点而不是左端点时的估算值.

均匀划分

取均匀划分总会是很方便的. 这表示, 每一个小区间都有相同的宽度, 并且要计算出其宽度也不是很难的事情. 如果积分区间是 $[a,\ b]$, 那么其长度是 $b-a$ 单位, 因此如果将该区间 n 等分, 那么每一个小区间的长度是 $(b-a)/n$ 单位. 我们称这个量为 h, 故 $h=(b-a)/n$. 此外, 出现在黎曼和定义中的表达式 $(x_j - x_{j-1})$ 正是第 j 个条纹的宽度, 因此它正是 h. 我们的表达式

$$\sum_{j=1}^{n} f(c_j)(x_j - x_{j-1})$$

可以简化为

$$h \times \sum_{j=1}^{n} f(c_j).$$



你仍然需要选取数 c_j, 但这一次就简单多了. 例如, 我们使用 10 个等宽的条纹来估算积分

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x,$$

每一条的宽度是 $h=(2-0)/10$, 即 1/5, 而且 $n=10$. 因此, 我们有

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx h \times \sum_{j=1}^{n} f(c_j) = \frac{1}{5}\sum_{j=1}^{10} \mathrm{e}^{-c_j^2}.$$

这些区间的宽度都是 1/5, 因此从 0 开始, 我们看到了如下的划分:

$$0 < \frac{1}{5} < \frac{2}{5} < \frac{3}{5} < \frac{4}{5} < 1 < \frac{6}{5} < \frac{7}{5} < \frac{8}{5} < \frac{9}{5} < 2.$$

如果我们令 c_j 为每一个小区间的右端点, 那么就有 $c_1 = \frac{1}{5}$, $c_2 = \frac{2}{5}$, 以此类推直到

$c_{10}=2$. 我们将这些数代入上述公式中, 得到

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx \frac{1}{5}\left(\mathrm{e}^{-(1/5)^2}+\mathrm{e}^{-(2/5)^2}+\cdots+\mathrm{e}^{-(9/5)^2}+\mathrm{e}^{-2^2}\right).$$

在这个和中有 10 项. 由于函数 f 在 0 和 2 之间是递减的, 而且我们使用了每一条的右端点, 因而以上就是估算过低的情况. (你知道为什么吗？) 不管怎样, 你可以使用计算器或计算机来求上面的和, 大约是 0.783 670(近似到小数点后六位).

如果使用每一个小区间的中点, 而不是左端点或右端点, 情况又会怎样呢？我们知道, $\left[0, \frac{1}{5}\right]$ 的中点是 $\frac{1}{10}$, $\left[\frac{1}{5}, \frac{2}{5}\right]$ 的中点是 $\frac{3}{10}$, 以此类推. 因此, 另一个可能的近似是

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx \frac{1}{5}(\mathrm{e}^{-(1/10)^2}+\mathrm{e}^{-(3/10)^2}+\cdots+\mathrm{e}^{-(17/10)^2}+\mathrm{e}^{-(19/10)^2}).$$

这大约是 0.882 202.

B.2　梯 形 法 则

涉及选取数 c_j 的问题是很困难的. 大多数情况下, 人们或者选择左端点或者选择右端点, 中点也是个常见的 (并且合理的) 选择. 这里还有一种估算积分的方法, 它不需要选择 (当然是在你决定使用这种方法的时候!) 但会给出更好的估算. 它被称作梯形法则.

其基本思想非常简单：我们允许条纹的上边不平行于底边. 每一条纹的上边都是连接曲线 $y=f(x)$ 上的两个相应点的线段. 图 B-3 就是说明这两种方法间区别的图像.

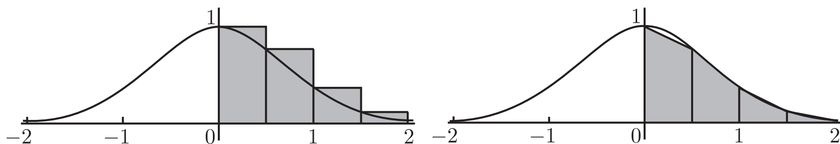


图　B-3

让我们来好好看看其中的一条新条纹, 如图 B-4 所示.

由于有两条边是平行的, 故该条纹是一个梯形. 底边长是 (x_j-x_{j-1}) 单位, 而平行的边的高度为 $f(x_{j-1})$ 单位和 $f(x_j)$ 单位. 根据梯形面积公式, 这个梯形条纹的面积是 $\frac{1}{2}(f(x_{j-1})+f(x_j))(x_j-x_{j-1})$ 平方单位. 如果我们确保划分都是均匀的, 那么如同上一节, 可知 x_j-x_{j-1} 就是 $(b-a)/n$. 这恰好就是条纹的宽度 (单位), 我们称之为 h, 因此, 一个条纹的面积变为

$$\frac{h}{2}(f(x_{j-1})+f(x_j))$$

平方单位. 余下的工作就是把所有的梯形条纹面积都加在一起. 我们可以只将一个 Σ 符号放在以上量的外面, 提取常数因子 $h/2$, 即

$$\int_a^b f(x)\mathrm{d}x \approx \frac{h}{2}\sum_{j=1}^{n}(f(x_{j-1})+f(x_j)).$$

事实上, 我们可以把这个表达式再简化一些. 你看, 除了最左边和最右边的条纹, 其他的相邻条纹都共用一条边, 如图 B-5 所示.

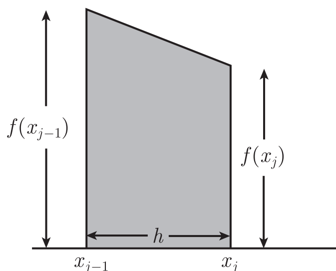


图 B-4

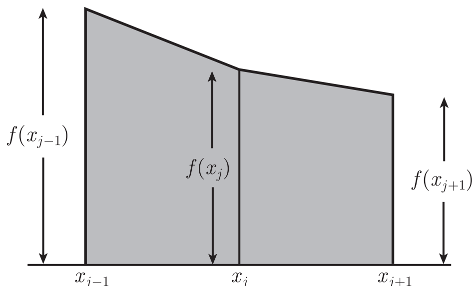


图 B-5

这意味着, 我们可以将很多项合并. 特别地, 除了 x_0 和 x_n 之外, 形如 $f(x_j)$ 的每一项都被用到两次. 例如 $n=4$ 时, 我们有

$$\int_a^b f(x)\mathrm{d}x \approx \frac{h}{2}((f(x_0)+f(x_1))+(f(x_1)+f(x_2))+f(x_2)+f(x_3))+(f(x_3)+f(x_4))).$$

因此, 我们可以将和式中除第一项和最后一项外的所有项合并, 得到

$$\int_a^b f(x)\mathrm{d}x \approx \frac{h}{2}(f(x_0)+2f(x_1)+2f(x_2)+2f(x_3)+f(x_4)).$$

同样的技巧适用于一般情况, 因此有

梯形法则：如果 $x_0<x_1<\cdots<x_n$ 是 $[a,\ b]$ 上的均匀划分, 且 $h=(b-a)/n$ 是条宽, 那么

$\int_a^b f(x)\mathrm{d}x \approx \frac{h}{2}(f(x_0)+2f(x_1)+2f(x_2)+\cdots+2f(x_{n-2})+2f(x_{n-1})+f(x_n)).$



让我们应用它来求下面积分的近似值：

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x.$$

我们取 $n=5$. 由于 $[0,\ 2]$ 的长度为 2 单位, 从而每一条的宽度为 $h=\frac{2}{5}$ 单位, 且划分是

$$0<\frac{2}{5}<\frac{4}{5}<\frac{6}{5}<\frac{8}{5}<2.$$

根据梯形法则, 我们有

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx \frac{2/5}{2}\left(\mathrm{e}^{-0^2}+2\mathrm{e}^{-(2/5)^2}+2\mathrm{e}^{-(4/5)^2}+2\mathrm{e}^{-(6/5)^2}+2\mathrm{e}^{-(8/5)^2}+\mathrm{e}^{-2^2}\right).$$

如果你愿意, 也可以将右边简化为

$$\frac{1}{5}\left(1+2\mathrm{e}^{-4/25}+2\mathrm{e}^{-16/25}+2\mathrm{e}^{-36/25}+2\mathrm{e}^{-64/25}+\mathrm{e}^{-4}\right).$$

你可以使用计算器或计算机来计算, 结果近似到小数点后六位是 0.881 131. 这比我们在 B.1 节结尾部分求出的估算 1.0865 略小一点, 但它很接近 B.1.1 节结尾部分的估算 0.882 202.

B.3　辛普森法则

为什么要止步于梯形法则呢? 梯形仍然有一个笨拙的线形上边. 在条纹的上边使用曲线而不是线段, 我们可以做得更好. 以下就是操作细节. 首先, 我们来看看相邻的两个条纹, 不用线段连接上边, 而是用一个二次曲线, 如图 B-6 所示.

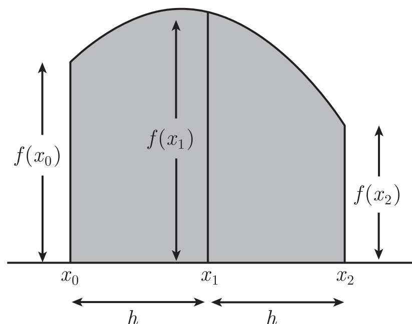


图　B-6

正如我们将在 B.3.1 节中看到的, 阴影部分的面积是

$$\frac{h}{3}(f(x_0)+4f(x_1)+f(x_2))$$

平方单位, 其中我们又设了 $h=(b-a)/n$. 现在, 如果对每一对条纹重复这个操作, 再将所有的面积相加, 就会得到近似. 如同梯形法则的情况, 相邻的两个条纹共用一条边, 因此会有一些量被重复一次. 例如, 如果有四个条纹, 那么面积和将是

$$\frac{h}{3}((f(x_0)+4f(x_1)+f(x_2))+(f(x_2)+4f(x_3)+f(x_4)));$$

我们把形如 $f(x_2)$ 的两项合并起来变为 $2f(x_2)$, 因此面积和是

$$\frac{h}{3}(f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+f(x_4)).$$

如果有更多的条纹依然会有相同样式的结果. 如果 j 是偶数, $f(x_j)$ 的系数等于 2; 如果 j 是奇数, $f(x_j)$ 的系数等于 4——$f(x_0)$ 和 $f(x_n)$ 除外, 它们的系数都是 1. 总之, 我们有:

> 辛普森法则: 如果 n 是偶数, $x_0<x_1<\cdots<x_n$ 是 $[a, b]$ 上的均匀划分, 且 $h=(b-a)/n$ 是条宽, 那么
> $\int_a^b f(x)\mathrm{d}x \approx \frac{h}{3}(f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+\cdots+2f(x_{n-2})+4f(x_{n-1})+f(x_n)).$

我们拿它和上一节的梯形法则比较一下. 代替形如 1, 2, 2, ..., 2, 2, 1 的系数, 这

一次系数形如 1, 4, 2, 4, 2, ..., 2, 4, 2, 4, 1. 还要注意的是, 前面分母中的常数为 3 而不是 2.

应用辛普森法则很容易. 我们回到原来的那个例子中:

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x,$$

并应用辛普森法则, 其中 $n=8$. (我们不能用 $n=5$, 因为 n 必须为偶数才能使用辛普森法则.) 每一条的宽度为 $h=(2-0)/8$ 单位, 即 $\frac{1}{4}$, 因此划分为

$$0<\frac{1}{4}<\frac{1}{2}<\frac{3}{4}<1<\frac{5}{4}<\frac{3}{2}<\frac{7}{4}<2.$$

根据以上公式, 我们有

$$\begin{aligned}\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx \frac{1/4}{3}(&\mathrm{e}^{-0^2}+4\mathrm{e}^{-(1/4)^2}+2\mathrm{e}^{-(1/2)^2}+4\mathrm{e}^{-(3/4)^2}+2\mathrm{e}^{-1^2}\\&+4\mathrm{e}^{-(5/4)^2}+2\mathrm{e}^{-(3/2)^2}+4\mathrm{e}^{-(7/4)^2}+\mathrm{e}^{-2^2}).\end{aligned}$$

使用计算器, 这大约是 0.882 066, 这十分接近我们在上一节的估算. 确切地说, 使用梯形法则时 (其中 $n=5$), 我们得到估算 0.881 131. 为了准确起见, 我使用了计算机程序, 得积分近似到小数点后六位的正确值是 0.882 081. 因此, 辛普森法则 ($n=8$) 比梯形法则 ($n=5$) 更好. 当然, 更公平的比较需在两种情况下都使用 $n=8$; 希望你来重复这种情况下的梯形法则的计算, 并和刚才相应的辛普森法则的估算结果进行比较.

辛普森法则的证明

让我们将图像平移, 以便中线位于 y 轴, 如图 B-7 所示.

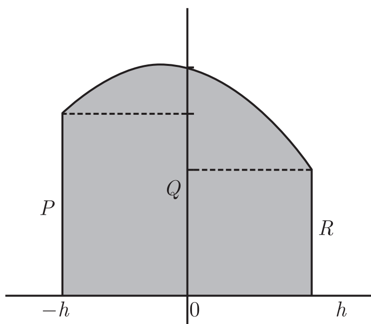


图 B-7

可以看到, 平移的结果将划分端点的 x 坐标移到了 $-h$、0 和 h. 不再使用 $f(x_0)$、$f(x_1)$ 和 $f(x_2)$, 我们只分别写出 P、Q 和 R. 上边的点由某二次曲线连接, 但我们不知道它是什么. 好吧, 我们就称它为 g 并假设 $g(x)=Ax^2+Bx+C$. 我们知道 $P=g(-h)$、$Q=g(0)$ 及 $R=g(h)$, 这表示

$$\begin{aligned}P&=A(-h)^2+B(-h)+C,\\Q&=A(0)^2+B(0)+C,\\R&=Ah^2+Bh+C.\end{aligned}$$

中间那个方程就是 $C=Q$, 那么重新整理其他两个方程, 会看到 $A=(P+R-2Q)/(2h^2)$. (我们不需要知道 B 是什么!) 现在, 所求阴影部分的面积简化后就是

$$\int_{-h}^{h}(Ax^2+Bx+C)\mathrm{d}x=\left(\frac{A}{3}x^3+\frac{B}{2}x^2+Cx\right)\Big|_{-h}^{h}=\frac{2Ah^3}{3}+2Ch$$

平方单位. 从上述公式中代换 A 和 C 的值, 表达式可简化为

$$\frac{2h^3}{3}\times\frac{P+R-2Q}{2h^2}+2Qh=\frac{h}{3}(P+4Q+R).$$

现在, 我们所要做的就是将它平移至更一般的位置 (不影响其面积) 并用函数值 $f(x_0)$、$f(x_1)$ 和 $f(x_2)$ 分别替换 P、Q 和 R, 来获得上一节开始部分的原型公式.

B.4 近似的误差

做近似 (或估算, 如果你更喜欢这个词) 的意义就是求接近于你要找的真实量的结果. 如果你真的能够确切地回答这个问题, 你就应该去做, 但有些时候这太难了. 因此, 近似至少可以给你提供接近于真实值的一个数. 正如我们多次看到的, 特别是当我们讨论线性化以及泰勒级数的时候 (见 13.2 节及 25.3 节), 还有一个重要的问题：近似有多好呢？你的近似是至少接近真实值, 还是在四周打转呢？

为了将这个问题量化, 我们再来看看近似中的误差, 它就是真实量和近似之间的差. 因此, 假设我们使用上述技巧中的一个 —— 均匀划分的条纹、梯形法则或辛普森法则 —— 来近似积分 $\int_a^b f(x)\,\mathrm{d}x$. 我们会得到

$$\int_a^b f(x)\mathrm{d}x\approx A,$$

其中 A 是近似值. 误差的绝对值是

$$|\text{误差}|=\left|\int_a^b f(x)\mathrm{d}x-A\right|.$$

事实表明, 通过 f 的导数 (如果它们存在), 我们可以对误差大小有些了解. 在那种情况下, 我们可以设 M_1 是 $|f'(x)|$ 在 $[a,\ b]$ 上的最大值. 类似地, 设 M_2 是 $|f''(x)|$ 在 $[a,\ b]$ 上的最大值, 最后设 M_4 是 $\left|f^{(4)}(x)\right|$ 在 $[a,\ b]$ 上的最大值. 那么, 我们可以证明下列误差的范围, 这取决于所使用的方法：

对于均匀划分的条纹, $\quad |\text{误差}|\leqslant\dfrac{1}{2}M_1(b-a)h,$

对于梯形法则, $\quad |\text{误差}|\leqslant\dfrac{1}{12}M_2(b-a)h^2,$

对于辛普森法则, $\quad |\text{误差}|\leqslant\dfrac{1}{180}M_4(b-a)h^4.$

如往常一样, 这里的 h 是条纹宽度 $(b-a)/n$. 尽管上述公式都很相似, 但是它们还是有所不同的. 首先, 前面的系数不一样. 其次, 所涉及的导数不同：对于条纹, 出现的是一阶导 (M_1 的形式); 对于梯形法则, 出现的是二阶导; 而对于辛普森法则,

则是四阶导. 然而, 最显著的区别是 h 的次数. 这显示了条纹宽度变小时, 误差减少的程度, 这当然发生在你取了很多条纹的时候. 当 h 变小时, h^4 会比 h^2 或 h 更快变小, 因此, 当使用很多条纹时, 辛普森法则与其他方法相较更胜一筹.

B.4.1 估算误差的例子



我们来看看这个附录中早早出现的例子

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x$$

中误差结果的情况, 首先, 我们设 $f(x)=\mathrm{e}^{-x^2}$, 然后计算

$$f'(x)=-2x\mathrm{e}^{-x^2},\quad f''(x)=(4x^2-2)\mathrm{e}^{-x^2},\quad f^{(3)}(x)=-4x(2x^2-3)\mathrm{e}^{-x^2},$$
$$f^{(4)}(x)=4(4x^4-12x^2+3)\mathrm{e}^{-x^2}.$$

首先, 我们来求 M_1. 这表示, 我们需要求出 $|f'(x)|$ 在 [0, 2] 上的最大值, 它实际上是 $-f'(x)$. 由于二阶导 $f''(x)$ 在 $x=1/\sqrt{2}$ 时为 0, 并且在那里其符号由负变为正, 故在 $1/\sqrt{2}$ 处 $f'(x)$ 有一个局部最小值. 这意味着, $f'(x)$ 在 [0, 2] 上的最小值是 $-\sqrt{2}e^{-1/2}$, 因此, $|f'(x)|$ 的最大值是 $\sqrt{2}\mathrm{e}^{-1/2}$. 即 $M_1=\sqrt{2}\mathrm{e}^{-1/2}$.

现在, 我们可以回到 B.1.1 节的积分估算中了. 那里, 我们使用了 10 个均匀划分的条纹来估算积分. 由于 $a=0$, $b=2$, $h=(2-0)/10=\frac{1}{5}$, 故我们有

$$|\text{使用 10 个均匀划分的条纹的误差}|\leqslant\frac{1}{2}M_1(b-a)h=\frac{1}{2}\times\sqrt{2\mathrm{e}}^{-1/2}(2-0)\frac{1}{5}$$
$$=\frac{\sqrt{2}}{5}\mathrm{e}^{-1/2}.$$

这大约是 0.171 553. 注意, 不管你使用左端点、右端点或中间的某个点作为 c_n, 都不要紧. (在 B.1.1 节, 我们使用了右端点和中点来求两个不同的估算, 它们都精确到大概 ±0.171 553.)

我们再来看看梯形法则. 在 B.2 节, 我们使用了 5 个宽度为 $h=2/5$ 的梯形来估算积分 (故 $n=5$). 为了查看误差会有多大, 我们需要在 [0, 2] 上最大化 $|f''(x)|$ 来求 M_2. 为此, 回头看看上述公式中的 $f^{(2)}(x)$ 和 $f^{(3)}(x)$. $f^{(3)}(x)$ 在 [0, 2] 上的零点在 $x=0$ 和 $x=\sqrt{3/2}$, 因此, 这些点就是 $f^{(2)}(x)$ 的临界点. (请记住, 三阶导是二阶导的导数!) 因此, 我们可以检验 $f''(0)$ 和 $f''(\sqrt{3/2})$ 的值, 还有在另一个端点 2 上的 $f''(2)$ 的值. 我们求出 $f''(0)=-2$, $f''(\sqrt{3/2})=4e^{-3/2}$, $f''(2)=14\mathrm{e}^{-4}$. 它们绝对值当中的最大值是 $f''(0)$. 这意味着 $M_2=2$. 现在, 我们可以估算误差了 (记住 $h=2/5$):

$$|\text{使用 5 个梯形的误差}|\leqslant\frac{1}{12}M_2(b-a)h^2=\frac{1}{12}\times 2(2-0)\left(\frac{2}{5}\right)^2=\frac{4}{75},$$

这大概是 0.053 333.... 这比使用 10 个条纹的误差要小很多, 尽管我们只使用了 5 个

梯形! 由于我们之前的估算大约是 0.881 131, 我们证明了

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx 0.881\,131$$

这个近似可精确到 ±0.053 333. (这当然和我们在 B.3 节结尾部分的观察是一致的, 其中, 近似到小数点后六位的正确值实际上是 0.882 081.)

最后, 我们使用辛普森法则来估算误差. 在 B.3 节, 我们使用了 $n=8$ 时的辛普森法则来证明

$$\int_0^2 \mathrm{e}^{-x^2}\mathrm{d}x \approx 0.882\,066.$$

我们需要求 M_4, 它是 $\left|f^{(4)}(x)\right|$ 在 $[0,\ 2]$ 上的最大值. 这可能非常繁杂, 因为 $f^{(4)}(x) = 4\left(4x^4-12x^2+3\right)\mathrm{e}^{-x^2}$. 我们来分别求这三个因子的最大值, 以代替求整个式子的最大值. 对于 4 没有任何问题, e^{-x^2} 是正的且在 $x=0$ 上达到最大 (其最大值为 1); 因此, 我们只需要求出 $\left|4x^4-12x^2+3\right|$ 在 $[0,\ 2]$ 上的最大值点. 我们有

$$\frac{\mathrm{d}}{\mathrm{d}x}(4x^4-12x^2+3) = 16x^3-24x = 8x(2x^2-3),$$

因此, 要找的最大值点只能出现在临界点 $x=0$ 和 $x=\sqrt{3/2}$, 或另一个端点 $x=2$ 上. 将这些数代入, 我们可以求出最大值 19 出现在 $x=2$, 这意味着在 $[0,\ 2]$ 上有

$$|4x^4-12x^2+3| \leqslant 19.$$

综上所述, 我们可以说

$$M_4 \leqslant 4\times 19\times 1 = 76.$$

(实际上 $M_4=12$, 但是你需要看看 f 的五阶导, 这就够了!) 现在, 终于可以使用我们的公式了 $(h=(2-0)/8=1/4)$:

$$\begin{aligned}|\text{使用 } n=8 \text{ 的辛普森法则的误差}| &\leqslant \frac{1}{180}M_4(b-a)h^4 \\ &\leqslant \frac{1}{180}\times 76(2-0)\left(\frac{1}{4}\right)^4 = \frac{19}{5760}.\end{aligned}$$

这大约是 0.003 299, 它比我们之前计算的那两个误差更小一些.

B.4.2 误差项不等式的证明

证明 B.4 节三个误差不等式的后两个, 有点超出本书的范围, 但是第一个并不难证明:

$$|\text{使用 } n \text{ 个均匀划分条纹宽度为 } h \text{ 的误差}| \leqslant \frac{1}{2}M_1(b-a)h,$$

其中, M_1 是 $f'(x)$ 在 $[a,\ b]$ 上的最大值. 假设我们使用左端点来做估算. 我们就来看看其中的一条吧. 如果它的底是区间 $[q,\ q+h]$(对于某个 q), 那么它看起来就如图 B-8 所示.

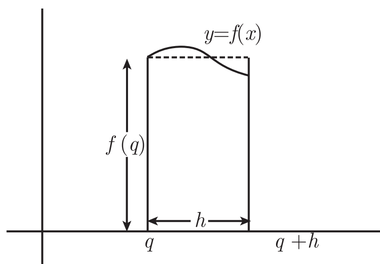


图 B-8

近似矩形的高度是 $f(q)$ 且宽度为 h 个单位, 因此, 近似的面积是 $hf(q)$ 平方单位. 一般地, 这个近似的结果会有多糟呢? 这完全取决于 f 的图像和常数直线 $y=f(q)$ 的偏离程度. 图 B-9 就是两种最坏的情况.

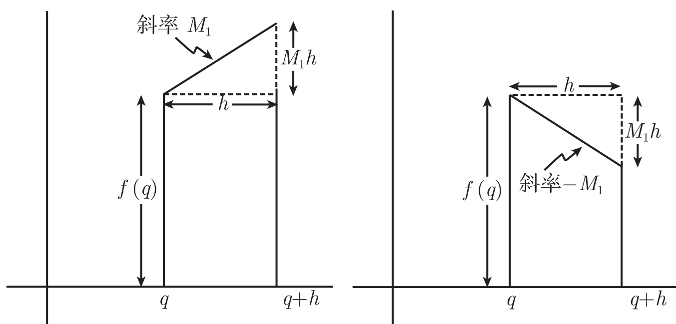


图 B-9

第一个图像显示了一条始于 $(q,\ f(q))$ 且斜率为 M_1 的线段, 而第二个图像显示了一条始于同一点且斜率为 $-M_1$ 的线段. 事实上, 该函数一定被夹在这两个极值之间. 的确, 第一条直线的方程为 $y=f(q)+M_1(x-q)$. 如果 $f(x)$ 高过这条线 (对于在区间 $[q,\ q+h]$ 内的 x), 那么我们有

$$f(x)>f(q)+M_1(x-q)\ \text{或}\ \frac{f(x)-f(q)}{x-q}>M_1.$$

根据中值定理 (见 11.3 节), 对于在 $[q,\ x]$ 上的某个 c, 左边部分等于 $f'(c)$, 故 $f'(c)>M_1$. 这是不可能的, 因为 M_1 是 $|f'(x)|$ 在 $[a,\ b]$ 上的最大值. 类似的论证显示了 $y=f(x)$ 总是位于那条向下倾的线的上方.

现在, 我们可以来看看误差了. 在第一种最坏的情况中, 真正的区域包括该条纹及一个边长为 h 与 M_1h 个单位的三角形; 在第二种最坏的情况中, 实际上从该条纹中去除了一个同样的三角形. 不管在哪一种情况中, 可能会偏离的面积是这个三角形的面积, 即 $\frac{1}{2}M_1h^2$ 平方单位. 剩下要做的就是, 用这个误差和条纹的个数 n 相乘, 会看到我们的近似不可能再比 $\frac{1}{2}M_1h^2n$ 更糟了. 事实上, 我们可以拿掉 h 的某个因子, 并使用等式 $nh=(b-a)$ 将上述表达式写作 $\frac{1}{2}M_1(b-a)h$. 这就是我们想要的了! 有时我们不是必须选择左端点的, 其他情形就请你来重复上述的证明. (事实上, 如果你使用中点, 可以证明误差实际上仅仅是 $\frac{1}{4}M_1(b-a)h$.)



符 号 列 表

符号	意义
$\mathbb{R}$	实数集合
$[a,b]$	从 a 到 b 的闭区间
(a,b)	从 a 到 b 的开区间
$(a,b]$	从 a 到 b 的半开区间
$A \setminus B$	在 A 中但不在 B 中的数
$f(x)$	以 x 为变量的函数
f^{-1}	函数 f 的反函数
$f \circ g$	函数 f 和 g 的复合函数
Δ	二次函数判别式
$\|x\|$	x 的绝对值
$\sin, \cos, \tan$	基本三角函数 (正弦、余弦、正切)
$\sec, \csc, \cot$	基本三角函数的倒数 (正割、余割、余切)
$\sin^{-1}, \cos^{-1}, \tan^{-1}$	反三角函数 (反正弦、反余弦、反正切)
$\sec^{-1}, \csc^{-1}, \cot^{-1}$	反三角函数的倒数 (反正割、反余割、反余切)
$\sinh, \cosh, \tanh$	基本的双曲函数 (双曲正弦、双曲余弦、双曲正切)
$\mathrm{sech}, \cosh, \coth$	基本双曲函数的倒 (双曲正割、双曲余割、双曲余切)
$\sinh^{-1}, \cosh^{-1}, \tanh^{-1}$	基本双曲函数的反三角函数 (反双曲正弦、反双曲余弦、反双曲正切)
$\mathrm{sech}^{-1}, \mathrm{csch}^{-1}, \coth^{-1}$	基本双曲函数的反三角函数的倒数 (反双曲正割、反双曲余割、反双曲余切)
$\ln(x), \log_{\mathrm{e}}(x)$	x 的自然对数
$\lim\limits_{x \to a}$	当 x 趋于 a 时的双方向极限
$\lim\limits_{x \to a^+}$	当 x 趋于 a 时的右极限
$\lim\limits_{x \to a^-}$	当 x 趋于 a 时的左极限
DNE	极限不存在
$0/0, \infty/\infty, 0 \times \infty$	不定式
$0^0, 1^\infty, \infty^0$	不定式
$\overset{\text{l'H}}{=\!=\!=}$	等于, 使用洛必达法则
$\sim$	渐近函数或数列
$\approx$	约等于
Δx	自变量 x 所发生的变化
$f'(x)$	函数 f 关于 x 的导数

$f''(x), f^{(2)}(x)$	函数 f 关于 x 的二次导数
$f^{(n)}(x)$	函数 f 关于 x 的 n 次导数
$\frac{\mathrm{d}y}{\mathrm{d}x}, \frac{\mathrm{d}}{\mathrm{d}x}(y), \mathrm{d}y/\mathrm{d}x$	y 关于 x 的导数
$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}, \frac{\mathrm{d}^2}{\mathrm{d}x^2}(y)$	y 关于 x 的二次导数
x, v, a	位移, 速度, 加速度
g	重力加速度
$\|AB\|$	线段 AB 的长度
ΔABC	以 A、B、C 为顶点的三角形
e	自然对数的底
$t_{1/2}$	放射性物质的半衰期
$\star$	不连续（使用在符号表格里）
$L(x)$	线性化
$\mathrm{d}f$	函数 f 的微分
$\sum_{j=a}^{b}$	从 $j=a$ 到 b 的和
$F(x)\Big\|_a^b$	$F(b)-F(a)$
$\int_a^b f(x)\mathrm{d}x$	函数 f 关于 x 的定积分
$\int f(x)\mathrm{d}x$	函数 f 关于 x 的不定积分 (反导数)
f_{av}	函数 f 的平均值
I_n	积分数 n (递归公式)
$\{a_n\}$	数列 $a_1, a_2, a_3\cdots$
$\sum_{n=1}^{\infty} a_n$	无穷级数 $a_1+a_2+a_3+\cdots$
$n!$	n 的阶乘 $(1\times 2\times 3\times\cdots\times(n-1)\times n)$
$P_N(x)$	N 阶泰勒多项式
$R_N(x)$	n 阶余项
(r,θ)	极坐标
i	$\sqrt{-1}$
$z=x+\mathrm{i}y$	笛卡儿形式的复数
$z=r\mathrm{e}^{\mathrm{i}\theta}$	极坐标形式的复数

e^z	以 z 为指数的复数
$\mathrm{Re}(z)$	z 的实部
$\mathrm{Im}(z)$	z 的虚部
$\bar{z}$	z 的复共轭
$\lvert z\rvert$	z 的模
$\arg(z)$	z 的辐角
y_H	齐次解 (微分方程)
y_P	特解 (微分方程)

索　引

H

J

K

L

M

N

O

P

Q

R

S

T

W

X

Y

Z

对于学习微积分有困难的同学来说，这是一本难能可贵的参考书。

——《数学教师》杂志

班纳的写作风格引人入胜，一点儿也不古板或令人生畏，他努力阐释解题的所有步骤。因其独到的讲解，本书成为了广大微积分教师的“得力助手”。

——《美国数学月刊》

本书语言平实，亲和力十足，是广大微积分学习者的良师益友。班纳的书写得非常到位，而且非常吸引读者。

—— Gerald B. Folland，《高等微积分》作者

普林斯顿微积分读本（修订版）

对于大多数学生来说，微积分或许是他们曾经上过的倍感迷茫且最受挫折的一门课程了。本书不仅让学生们能有效地学习微积分，更重要的是提供了战胜微积分的必备工具。

这本经典著作源于风靡美国普林斯顿大学的阿德里安·班纳教授的微积分复习课程，将易用性与可读性以及内容的深度与数学的严谨完美地结合在了一起，激励学生不再惧怕微积分，并在考试中获得高分。

作者班纳是美国普林斯顿大学的著名数学教授，并担任新技术研究中心主任。他的授课风格非正式、有吸引力并完全不强求，甚至在不失其详尽性的基础上又增添了许多娱乐性，而且他不会跳过讨论一个问题的任何步骤。



图灵社区：iTuring.cn
热线：(010) 51095186 转 600

分类建议 数学 / 微积分

人民邮电出版社网址：www.ptpress.com.cn

ISBN 978-7-115-43559-0
定价：99.00元

看完了

如果您对本书内容有疑问，可访问图灵社区www.ituring.com.cn ，联系负责编辑或作译者，参与本书讨论。

如果是有关电子书的建议或问题，请联系专用客服邮箱:
ebook@turingbook.com。

在这可以找到我们:

微博 @图灵教育：好书、活动每日播报
微博 @图灵新知：图灵教育的科普小组
微信 图灵教育：turingbooks